

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

**EP 1 310 571 A2**

(12)

**EUROPEAN PATENT APPLICATION**

(43) Date of publication:

**14.05.2003 Bulletin 2003/20**

(51) Int Cl.7: **C12Q 1/70**, C07K 14/015,  
C12N 15/86, C12N 15/10,  
C12Q 1/68, C12N 5/10

(21) Application number: **02257826.4**

(22) Date of filing: **12.11.2002**

(84) Designated Contracting States:

**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR  
IE IT LI LU MC NL PT SE SK TR**

Designated Extension States:

**AL LT LV MK RO SI**

(30) Priority: **13.11.2001 US 350607 P**

**17.12.2001 US 341117 P**

**01.05.2002 US 377066 P**

**05.06.2002 US 386675 P**

(71) Applicant: **The Trustees of The University of  
Pennsylvania**

**Philadelphia, Pennsylvania 19104-6283 (US)**

(72) Inventors:

- **Gao, Guangping**  
**Rosemont, Pennsylvania 19010 (US)**

- **Wilson, James M.**  
**Gladwyne, Pennsylvania 19035 (US)**

- **Alvira, Maricio**  
**Philadelphia, Pennsylvania 19104 (US)**

(74) Representative: **Hale, Stephen Geoffrey**

**JY & GW Johnson,  
Kingsbourne House,  
229-231 High Holborn  
London WC1V 7DP (GB)**

(54) **A Method of detecting and/or identifying adeno-associated virus (AAV) sequences and isolating novel sequences identified thereby**

(57) A method for detecting and isolating AAV sequences in a sample of DNA obtained from tissue or cells is provided. The invention further provides AAV sequences identified by this method, and vectors constructed using these sequences.

**EP 1 310 571 A2**

**Description****BACKGROUND OF THE INVENTION**

**[0001]** Adeno-associated virus (AAV), a member of the Parvovirus family, is a small nonenveloped, icosahedral virus with single-stranded linear DNA genomes of 4.7 kilobases (kb) to 6 kb. AAV is assigned to the genus, Dependovirus, because the virus was discovered as a contaminant in purified adenovirus stocks. AAV's life cycle includes a latent phase at which AAV genomes, after infection, are site specifically integrated into host chromosomes and an infectious phase in which, following either adenovirus or herpes simplex virus infection, the integrated genomes are subsequently rescued, replicated, and packaged into infectious viruses. The properties of non-pathogenicity, broad host range of infectivity, including non-dividing cells, and potential site-specific chromosomal integration make AAV an attractive tool for gene transfer.

**[0002]** Recent studies suggest that AAV vectors may be the preferred vehicle for gene therapy. To date, there have been 6 different serotypes of AAVs isolated from human or non-human primates (NHP) and well characterized. Among them, human serotype 2 is the first AAV that was developed as a gene transfer vector; it has been widely used for efficient gene transfer experiments in different target tissues and animal models. Clinical trials of the experimental application of AAV2 based vectors to some human disease models are in progress, and include such diseases as cystic fibrosis and hemophilia B.

**[0003]** What are desirable are AAV-based constructs for gene delivery.

**SUMMARY OF THE INVENTION**

**[0004]** In one aspect, the invention provides a novel method of detecting and identifying AAV sequences from cellular DNAs of various human and non-human primate (NHP) tissues using bioinformatics analysis, PCR based gene amplification and cloning technology, based on the nature of latency and integration of AAVs in the absence of helper virus co-infection.

**[0005]** In another aspect, the invention provides method of isolating novel AAV sequences detected using the above described method of the invention. The invention further comprises methods of generating vectors based upon these novel AAV serotypes, for serology and gene transfer studies solely based on availability of capsid gene sequences and structure of rep/cap gene junctions.

**[0006]** In still another aspect, the invention provides a novel method for performing studies of serology, epidemiology, biodistribution and mode of transmission, using reagents according to the invention, which include generic sets of primers/probes and quantitative real time PCR.

**[0007]** In yet another aspect, the invention provides a method of isolating complete and infectious genomes of novel AAV serotypes from cellular DNA of different origins using RACE and other molecular techniques.

**[0008]** In a further aspect, the invention provides a method of rescuing novel serotypes of AAV genomes from human and NHP cell lines using adenovirus helpers of different origins.

**[0009]** In still a further aspect, the invention provides novel AAV serotypes, vectors containing same, and methods of using same.

**[0010]** These and other aspects of the invention will be readily apparent from the following detailed description of the invention.

**BRIEF DESCRIPTION OF THE DRAWINGS**

**[0011]**

Figs. 1A through 1AAAR provide an alignment of the nucleic acid sequences encoding at least the cap proteins for the AAV serotypes. The full-length sequences including the ITRs, the rep region, and the capsid region are provided for novel AAV serotype 7 [SEQ ID NO:1], and for previously published AAV1 [SEQ IN NO:6], AAV2 [SEQ ID NO:7]; and AAV3 [SEQ ID NO:8] Novel AAV serotypes AAV8 [SEQ ID NO:4] and AAV9 [SEQ ID NO:5] are the subject of co-filed applications. The other novel clones of the invention provided in this alignment include: 42-2 [SEQ ID NO:9], 42-8 [SEQ ID NO:27], 42-15 [SEQ ID NO:28], 42-5b [SEQ ID NO: 29], 42-1b [SEQ ID NO:30]; 42-13 [SEQ ID NO: 31], 42-3a [SEQ ID NO: 32], 42-4 [SEQ ID NO:33], 42-5a [SEQ ID NO: 34], 42-10 [SEQ ID NO:35], 42-3b [SEQ ID NO: 36], 42-11 [SEQ ID NO: 37], 42-6b [SEQ ID NO:38], 43-1 [SEQ ID NO: 39], 43-5 [SEQ ID NO: 40], 43-12 [SEQ ID NO:41], 43-20 [SEQ ID NO:42], 43-21 [SEQ ID NO: 43], 43-23 [SEQ ID NO:44], 43-25 [SEQ ID NO: 45], 44.1 [SEQ ID NO:47], 44.5 [SEQ ID NO:47], 223.10 [SEQ ID NO:48], 223.2 [SEQ ID NO:49], 223.4 [SEQ ID NO:50], 223.5 [SEQ ID NO: 51], 223.6 [SEQ ID NO: 52], 223.7 [SEQ ID NO: 53], A3.4 [SEQ ID NO: 54], A3.5 [SEQ ID NO:55], A3.7 [SEQ ID NO: 56], A3.3 [SEQ ID NO:57], 42.12 [SEQ ID NO: 58], 44.2 [SEQ

ID NO: 59]. The nucleotide sequences of the signature regions of AAV10 [SEQ ID NO: 117], AAV11 [SEQ ID NO: 118] and AAV12 [SEQ ID NO: 119] are provided in this figure. Critical landmarks in the structures of AAV genomes are shown. Gaps are demonstrated by dots. The 3' ITR of AAV1 [SEQ ID NO: 6] is shown in the same configuration as in the published sequences. TRS represents terminal resolution site. Notice that AAV7 is the only AAV reported that uses GTG as the initiation codon for VP3.

Figs. 2A through 2F are an alignment of the amino acid sequences of the proteins of the vp1 capsid proteins of previously published AAV serotypes 1 [SEQ ID NO: 64], AAV2 [SEQ ID NO: 70], AAV3 [SEQ ID NO: 71], AAV4 [SEQ ID NO: 63], AAV5 [SEQ ID NO: 114], and AAV6 [SEQ ID NO: 65] and novel AAV sequences of the invention, including: C1 [SEQ ID NO: 60], C2 [SEQ ID NO: 61], C5 [SEQ ID NO: 62], A3-3 [SEQ ID NO: 66], A3-7 [SEQ ID NO: 67], A3-4 [SEQ ID NO: 68], A3-5 [SEQ ID NO: 69], 3.3b [SEQ ID NO: 62], 223.4 [SEQ ID NO: 73], 223-5 [SEQ ID NO: 74], 223-10 [SEQ ID NO: 75], 223-2 [SEQ ID NO: 76], 223-7 [SEQ ID NO: 77], 223-6 [SEQ ID NO: 78], 44-1 [SEQ ID NO: 79], 44-5 [SEQ ID NO: 80], 44-2 [SEQ ID NO: 81], 42-15 [SEQ ID NO: 84], 42-8 [SEQ ID NO: 85], 42-13 [SEQ ID NO: 86], 42-3A [SEQ ID NO: 87], 42-4 [SEQ ID NO: 88], 42-5A [SEQ ID NO: 89], 42-1B [SEQ ID NO: 90], 42-5B [SEQ ID NO: 91], 43-1 [SEQ ID NO: 92], 43-12 [SEQ ID NO: 93], 43-5 [SEQ ID NO: 94], 43-21 [SEQ ID NO: 96], 43-25 [SEQ ID NO: 97], 43-20 [SEQ ID NO: 99], 24.1 [SEQ ID NO: 101], 42.2 [SEQ ID NO: 102], 7.2 [SEQ ID NO: 103], 27.3 [SEQ ID NO: 104], 16.3 [SEQ ID NO: 105], 42.10 [SEQ ID NO: 106], 42-3B [SEQ ID NO: 107], 42-11 [SEQ ID NO: 108], F1 [SEQ ID NO: 109], F5 [SEQ ID NO: 110], F3 [SEQ ID NO: 111], 42-6B [SEQ ID NO: 112], 42-12 [SEQ ID NO: 113]. Novel serotypes AAV8 [SEQ ID NO: 95] and AAV9 [SEQ ID NO: 100] are the subject of co-filed patent applications.

Figs. 3A through 3C provide the amino acid sequences of the AAV7 rep proteins [SEQ ID NO: 3].

#### DETAILED DESCRIPTION OF THE INVENTION

**[0012]** In the present invention, the inventors have found a method which takes advantage of the ability of adeno-associated virus (AAV) to penetrate the nucleus, and, in the absence of a helper virus co-infection, to integrate into cellular DNA and establish a latent infection. This method utilizes a polymerase chain reaction (PCR)-based strategy for detection, identification and/or isolation of sequences of AAVs from DNAs from tissues of human and non-human primate origin as well as from other sources. Advantageously, this method is also suitable for detection, identification and/or isolation of other integrated viral and non-viral sequences, as described below.

**[0013]** The invention further provides nucleic acid sequences identified according to the methods of the invention. One such adeno-associated virus is of a novel serotype, termed herein serotype 7 (AAV7). Other novel adeno-associated virus serotypes provided herein include AAV10, AAV11, and AAV12. Still other novel AAV serotypes identified according to the methods of the invention are provided in the present specification. See, Figures and Sequence Listing, which is incorporated by reference.

**[0014]** Also provided are fragments of these AAV sequences. Among particularly desirable AAV fragments are the cap proteins, including the vp1, vp2, vp3, the hypervariable regions, the rep proteins, including rep 78, rep 68, rep 52, and rep 40, and the sequences encoding these proteins. Each of these fragments may be readily utilized in a variety of vector systems and host cells. Such fragments may be used alone, in combination with other AAV sequences or fragments, or in combination with elements from other AAV or non-AAV viral sequences. In one particularly desirable embodiment, a vector contains the AAV cap and/or rep sequences of the invention.

**[0015]** As described herein, alignments are performed using any of a variety of publicly or commercially available Multiple Sequence Alignment Programs, such as "Clustal W", accessible through Web Servers on the internet. Alternatively, Vector NTI utilities are also used. There are also a number of algorithms known in the art which can be used to measure nucleotide sequence identity, including those contained in the programs described above. As another example, polynucleotide sequences can be compared using Fasta, a program in GCG Version 6.1. Fasta provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences. For instance, percent sequence identity between nucleic acid sequences can be determined using Fasta with its default parameters (a word size of 6 and the NOPAM factor for the scoring matrix) as provided in GCG Version 6.1, herein incorporated by reference. Similar programs are available for amino acid sequences, e.g., the "Clustal X" program. Generally, any of these programs are used at default settings, although one of skill in the art can alter these settings as needed. Alternatively, one of skill in the art can utilize another algorithm or computer program which provides at least the level of identity or alignment as that provided by the referenced algorithms and programs.

**[0016]** The term "substantial homology" or "substantial similarity," when referring to a nucleic acid, or fragment thereof, indicates that, when optimally aligned with appropriate nucleotide insertions or deletions with another nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 95 to 99% of the aligned sequences. Preferably, the homology is over full-length sequence, or an open reading frame thereof, or another suitable fragment which is at least 15 nucleotides in length. Examples of suitable fragments are described herein.

**[0017]** The term "substantial homology" or "substantial similarity," when referring to amino acids or fragments thereof,

indicates that, when optimally aligned with appropriate amino acid insertions or deletions with another amino acid, there is amino acid sequence identity in at least about 95 to 99% of the aligned sequences. Preferably, the homology is over full-length sequence, or a protein thereof, e.g., a cap protein, a rep protein, or a fragment thereof which is at least 8 amino acids, or more desirably, at least 15 amino acids in length. Examples of suitable fragments are described herein.

**[0018]** By the term "highly conserved" is meant at least 80% identity, preferably at least 90% identity, and more preferably, over 97% identity. Identity is readily determined by one of skill in the art by resort to algorithms and computer programs known by those of skill in the art.

**[0019]** The term "percent sequence identity" or "identical" in the context of nucleic acid sequences refers to the residues in the two sequences which are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over the full-length of the genome, the full-length of a gene coding sequence, or a fragment of at least about 500 to 5000 nucleotides, is desired. However, identity among smaller fragments, e.g. of at least about nine nucleotides, usually at least about 20 to 24 nucleotides, at least about 28 to 32 nucleotides, at least about 36 or more nucleotides, may also be desired. Similarly, "percent sequence identity" may be readily determined for amino acid sequences, over the full-length of a protein, or a fragment thereof. Suitably, a fragment is at least about 8 amino acids in length, and may be up to about 700 amino acids. Examples of suitable fragments are described herein.

**[0020]** The AAV sequences and fragments thereof are useful in production of rAAV, and are also useful as antisense delivery vectors, gene therapy vectors, or vaccine vectors. The invention further provides nucleic acid molecules, gene delivery vectors, and host cells which contain the AAV sequences of the invention.

**[0021]** As described herein, the vectors of the invention containing the AAV capsid proteins of the invention are particularly well suited for use in applications in which the neutralizing antibodies diminish the effectiveness of other AAV serotype based vectors, as well as other viral vectors. The rAAV vectors of the invention are particularly advantageous in rAAV readministration and repeat gene therapy.

**[0022]** These and other embodiments and advantages of the invention are described in more detail below. As used throughout this specification and the claims, the terms "comprising" and "including" and their variants are inclusive of other components, elements, integers, steps and the like. Conversely, the term "consisting" and its variants is exclusive of other components, elements, integers, steps and the like.

## I. Methods of the Invention

### A. Detection of Sequences Via Molecular Cloning

**[0023]** In one aspect, the invention provides a method of detecting and/or identifying target nucleic acid sequences in a sample. This method is particularly well suited for detection of viral sequences which are integrated into the chromosome of a cell, e.g., adeno-associated viruses (AAV) and retroviruses, among others. The specification makes reference to AAV, which is exemplified herein. However, based on this information, one of skill in the art may readily perform the methods of the invention on retroviruses [e.g., feline leukemia virus (FeLV), HTLV I and HTLV II], and lentivirinae [e.g., human immunodeficiency virus (HIV), simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV), equine infectious anemia virus, and spumavirinal)], among others. Further, the method of the invention may also be used for detection of other viral and non-viral sequences, whether integrated or non-integrated into the genome of the host cell.

**[0024]** As used herein, a sample is any source containing nucleic acids, e.g., tissue, tissue culture, cells, cell culture, and biological fluids including, without limitation, urine and blood. These nucleic acid sequences may be DNA or RNA from plasmids, natural DNA or RNA from any source, including bacteria, yeast, viruses, and higher organisms such as plants or animals. DNA or RNA is extracted from the sample by a variety of techniques known to those of skill in the art, such as those described by Sambrook, Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory). The origin of the sample and the method by which the nucleic acids are obtained for application of the method of the invention is not a limitation of the present invention. Optionally, the method of the invention can be performed directly on the source of DNA, or on nucleic acids obtained (e.g., extracted) from a source.

**[0025]** The method of the invention involves subjecting a sample containing DNA to amplification via polymerase chain reaction (PCR) using a first set of primers specific for a first region of double-stranded nucleic acid sequences, thereby obtaining amplified sequences.

**[0026]** As used herein, each of the "regions" is predetermined based upon the alignment of the nucleic acid sequences of at least two serotypes (e.g., AAV) or strains (e.g., lentiviruses), and wherein each of said regions is composed of sequences having a 5' end which is highly conserved, a middle which is preferably, but necessarily, variable, and a 3' end which is highly conserved, each of these being conserved or variable relative to the sequences of the at least two aligned AAV serotypes. Preferably, the 5' and/or 3' end is highly conserved over at least about 9, and more preferably, at least 18 base pairs (bp). However, one or both of the sequences at the 5' or 3' end may be conserved over

more than 18 bp, more than 25 bp, more than 30 bp, or more than 50 bp at the 5' end. With respect to the variable region, there is no requirement for conserved sequences, these sequences may be relatively conserved, or may have less than 90, 80, or 70% identity among the aligned serotypes or strains.

**[0027]** Each of the regions may span about 100 bp to about 10 kilobase pairs in length. However, it is particularly desirable that one of the regions is a "signature region", i.e., a region which is sufficiently unique to positively identify the amplified sequence as being from the target source. For example, in one embodiment, the first region is about 250 bp in length, and is sufficiently unique among known AAV sequences, that it positively identifies the amplified region as being of AAV origin. Further, the variable sequences within this region are sufficiently unique that can be used to identify the serotype from which the amplified sequences originate. Once amplified (and thereby detected), the sequences can be identified by performing conventional restriction digestion and comparison to restriction digestion patterns for this region in any of AAV1, AAV2, AAV3, AAV4, AAV5, or AAV6, or that of AAV7, AAV10, AAV11, AAV12, or any of the other novel serotypes identified by the invention, which is predetermined and provided by the present invention.

**[0028]** Given the guidance provided herein, one of skill in the art can readily identify such regions among other integrated viruses to permit ready detection and identification of these sequences. Thereafter, an optimal set of generic primers located within the highly conserved ends can be designed and tested for efficient amplification of the selected region from samples. This aspect of the invention is readily adapted to a diagnostic kit for detecting the presence of the target sequence (e.g., AAV) and for identifying the AAV serotype, using standards which include the restriction patterns for the AAV serotypes described herein or isolated using the techniques described herein. For example, quick identification or molecular serotyping of PCR products can be accomplished by digesting the PCR products and comparing restriction patterns.

**[0029]** Thus, in one embodiment, the "signature region" for AAV spans about bp 2800 to about 3200 of AAV 1 [SEQ ID NO:6], and corresponding base pairs in AAV 2, AAV3, AAV4, AAV5, and AAV6. More desirably, the region is about 250 bp, located within bp 2886 to about 3143 bp of AAV 1 [SEQ ID NO:6], and corresponding base pairs in AAV 2 [SEQ ID NO:7], AAV3 [SEQ ID NO:8], and other AAV serotypes. See, Fig. 1. To permit rapid detection of AAV in the sample, primers which specifically amplify this signature region are utilized. However, the present invention is not limited to the exact sequences identified herein for the AAV signature region, as one of skill in the art may readily alter this region to encompass a shorter fragment, or a larger fragment of this signature region.

**[0030]** The PCR primers are generated using techniques known to those of skill in the art. Each of the PCR primer sets is composed of a 5' primer and a 3' primer. See, e.g., Sambrook et al, cited herein. The term "primer" refers to an oligonucleotide which acts as a point of initiation of synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced. The primer is preferably single stranded. However, if a double stranded primer is utilized, it is treated to separate its strands before being used to prepare extension products. The primers may be about 15 to 25 or more nucleotides, and preferably at least 18 nucleotides. However, for certain applications shorter nucleotides, e.g., 7 to 15 nucleotides are utilized.

**[0031]** The primers are selected to be sufficiently complementary to the different strands of each specific sequence to be amplified to hybridize with their respective strands. Therefore, the primer sequence need not reflect the exact sequence of the region being amplified. For example, a non-complementary nucleotide fragment may be attached to the 5' end of the primer, with the remainder of the primer sequence being completely complementary to the strand. Alternatively, non-complementary bases or longer sequences can be interspersed into the primer, provided that the primer sequence has sufficient complementarity with the sequence of the strand to be amplified to hybridize therewith and form a template for synthesis of the extension product of the other primer.

**[0032]** The PCR primers for the signature region according to the invention are based upon the highly conserved sequences of two or more aligned sequences (e.g., two or more AAV serotypes). The primers can accommodate less than exact identity among the two or more aligned AAV serotypes at the 5' end or in the middle. However, the sequences at the 3' end of the primers correspond to a region of two or more aligned AAV serotypes in which there is exact identity over at least five, preferably, over at least nine base pairs, and more preferably, over at least 18 base pairs at the 3' end of the primers. Thus, the 3' end of the primers is composed of sequences with 100% identity to the aligned sequences over at least five nucleotides. However, one can optionally utilize one, two, or more degenerate nucleotides at the 3' end of the primer.

**[0033]** For example, the primer set for the signature region of AAV was designed based upon a unique region within the AAV capsid, as follows. The 5' primer was based upon nt 2867-2891 of AAV2 [SEQ ID NO:7], 5'-GGTAATTCCTCCGGAATTGGCATT3'. See, Fig. 1. The 3' primer was designed based upon nt 3096-3122 of AAV2 [SEQ ID NO:7], 5'-GACTCATCAACAACAACTGGGGATTTC-3'. However, one of skill in the art may have readily designed the primer set based upon the corresponding regions of AAV1, AAV3, AAV4, AAV5, AAV6, or based upon the information provided herein, AAV7, AAV10, AAV11, AAV12, or another novel AAV of the invention. In addition, still other primer sets can be readily designed to amplify this signature region, using techniques known to those of skill in the art.

## B. Isolation of Target Sequences

**[0034]** As described herein, the present invention provides a first primer set which specifically amplifies the signature region of the target sequence, e.g., an AAV serotype, in order to permit detection of the target. In a situation in which further sequences are desired, e.g., if a novel AAV serotype is identified, the signature region may be extended. Thus, the invention may further utilize one or more additional primer sets.

**[0035]** Suitably, these primer sets are designed to include either the 5' or 3' primer of the first primer set and a second primer unique to the primer set, such that the primer set amplifies a region 5' or 3' to the signature region which anneals to either the 5' end or the 3' end of the signature region. For example, a first primer set is composed of a 5' primer, P1 and a 3' primer P2 to amplify the signature region. In order to extend the signature region on its 3' end, a second primer set is composed of primer P1 and a 3' primer P4, which amplifies the signature region and contiguous sequences downstream of the signature region. In order to extend the signature region on its 5' end, a third primer set is composed of a 5' primer, P5, and primer P2, such that the signature region and contiguous sequences upstream of the signature region are amplified. These extension steps are repeated (or performed at the same time), as needed or desired. Thereafter, the products results from these amplification steps are fused using conventional steps to produce an isolated sequence of the desired length.

**[0036]** The second and third primer sets are designed, as with the primer set for the signature region, to amplify a region having highly conserved sequences among the aligned sequences. Reference herein to the term "second" or "third" primer set is for each of discussion only, and without regard to the order in which these primers are added to the reaction mixture, or used for amplification. The region amplified by the second primer set is selected so that upon amplification it anneals at its 5' end to the 3' end of the signature region. Similarly, the region amplified by the third primer set is selected so that upon amplification it anneals at its 3' end anneals to the 5' end of the signature region. Additional primer sets can be designed such that the regions which they amplify anneal to the either the 5' end or the 3' end of the extension products formed by the second or third primer sets, or by subsequent primer sets.

**[0037]** For example, where AAV is the target sequence, a first set of primers (P1 and P2) are used to amplify the signature region from the sample. In one desirable embodiment, this signature region is located within the AAV capsid. A second set of primers (P1 and P4) is used to extend the 3' end of the signature region to a location in the AAV sequence which is just before the AAV 3' ITR, i.e., providing an extension product containing the entire 3' end of the AAV capsid when using the signature region as an anchor. In one embodiment, the P4 primer corresponds to nt 4435 to 4462 of AAV2 [SEQ ID NO:7], and corresponding sequences in the other AAV serotypes. This results in amplification of a region of about 1.6 kb, which contains the 0.25 kb signature region. A third set of primers (P3 and P2) is used to extend the 5' end of signature region to a location in the AAV sequences which is in the 3' end of the rep genes, i.e., providing an extension product containing the entire 5' end of the AAV capsid when using the signature region as an anchor. In one embodiment, the P3 primer corresponds to nt 1384 to 1409 of AAV2 [SEQ ID NO:7], and corresponding sequences in the other AAV serotypes. This results in amplification of a region of about 1.7 kb, which contains the 0.25 kb signature region. Optionally, a fourth set of primers are used to further extend the extension product containing the entire 5' end of the AAV capsid to also include the rep sequences. In one embodiment, the primer designated P5 corresponds to nt 108 to 133 of AAV2 [SEQ ID NO:7], and corresponding sequences in the other AAV serotypes and is used in conjunction with the P2 primer.

**[0038]** Following completion of the desired number of extension steps, the various extension products are fused, making use of the signature region as an anchor or marker, to construct an intact sequence. In the example provided herein, AAV sequences containing, at a minimum, an intact AAV cap gene are obtained. Larger sequences may be obtained, depending upon the number of extension steps performed.

**[0039]** Suitably, the extension products are assembled into an intact AAV sequence using methods known to those of skill in the art. For example, the extension products may be digested with DraIII, which cleaves at the DraIII site located within the signature region, to provide restriction fragments which are re-ligated to provide products containing (at a minimum) an intact AAV cap gene. However, other suitable techniques for assembling the extension products into an intact sequence may be utilized. See, generally, Sambrook et al, cited herein.

**[0040]** As an alternative to the multiple extension steps described above, another embodiment of the invention provides for direct amplification of a 3.1 kb fragment which allows isolation of full-length cap sequences. To directly amplify a 3.1 kb full-length cap fragment from NHP tissue and blood DNAs, two other highly conserved regions were identified in AAV genomes for use in PCR amplification of large fragments. A primer within a conserved region located in the middle of the rep gene is utilized (AV1ns: 5' GCTGCGTCAACTGGACCAATGAGAAC 3', nt of SEQ ID NO:6) in combination with the 3' primer located in another conserved region downstream of the Cap gene (AV2cas: 5' CGCAGAGACCAAAGTTCAACTGAAACGA 3', SEQ ID NO: 7) for amplification of AAV sequences including the full-length AAV cap. Typically, following amplification, the products are cloned and sequence analysis is performed with an accuracy of  $\geq 99.9\%$ . Using this method, the inventors have isolated at least 50 capsid clones which have subsequently been characterized. Among them, 37 clones were derived from Rhesus macaque tissues (rh.1 - rh.37), 6 clones from cy-

nomologous macaques (cy.1 - cy.6), 2 clones from Baboons (bb.1 and bb.2) and 5 clones from Chimps (ch.1 - ch.5). These clones are identified elsewhere in the specification, together with the species of animal from which they were identified and the tissues in that animal these novel sequences have been located.

#### C. Alternative method for isolating novel AAV

**[0041]** In another aspect, the invention provides an alternative method for isolating novel AAV from a cell. This method involves infecting the cell with a vector which provides helper functions to the AAV; isolating infectious clones containing AAV; sequencing the isolated AAV; and comparing the sequences of the isolated AAV to known AAV serotypes, whereby differences in the sequences of the isolated AAV and known AAV serotypes indicates the presence of a novel AAV.

**[0042]** In one embodiment, the vector providing helper functions provides essential adenovirus functions, including, e.g., E1a, E1b, E2a, E4ORF6. In one embodiment, the helper functions are provided by an adenovirus. The adenovirus may be a wild-type adenovirus, and may be of human or non-human origin, preferably non-human primate (NHP) origin. The DNA sequences of a number of adenovirus types are available from Genbank, including type Ad5 [Genbank Accession No. M73260]. The adenovirus sequences may be obtained from any known adenovirus serotype, such as serotypes 2, 3, 4, 7, 12 and 40, and further including any of the presently identified human types [see, e.g., Horwitz, cited above]. Similarly adenoviruses known to infect non-human animals (e.g., chimpanzees) may also be employed in the vector constructs of this invention. See, e.g., US Patent No. 6,083,716. In addition to wild-type adenoviruses, recombinant viruses or non-viral vectors (e.g., plasmids, episomes, etc.) carrying the necessary helper functions may be utilized. Such recombinant viruses are known in the art and may be prepared according to published techniques. See, e.g., US Patent No. 5,871,982 and US Patent 6,251,677, which describe a hybrid Ad/AAV virus. The selection of the adenovirus type is not anticipated to limit the following invention. A variety of adenovirus strains are available from the American Type Culture Collection, Manassas, Virginia, or available by request from a variety of commercial and institutional sources. Further, the sequences of many such strains are available from a variety of databases including, e.g., PubMed and GenBank.

**[0043]** In another alternative, infectious AAV may be isolated using genome walking technology (Siebert *et al.*, 1995, *Nucleic Acid Research*, **23**:1087-1088, Friezner-Degen *et al.*, 1986, *J. Biol. Chem.* **261**:6972-6985, BD Biosciences Clontech, Palo Alto, CA). Genome walking is particularly well suited for identifying and isolating the sequences adjacent to the novel sequences identified according to the method of the invention. For example, this technique may be useful for isolating inverted terminal repeat (ITRs) of the novel AAV serotype, based upon the novel AAV capsid and/or rep sequences identified using the methods of the invention. This technique is also useful for isolating sequences adjacent to other AAV and non-AAV sequences identified and isolated according to the present invention. See, Examples 3 and 4.

**[0044]** The methods of the invention may be readily used for a variety of epidemiology studies, studies of biodistribution, monitoring of gene therapy via AAV vectors and vector derived from other integrated viruses. Thus, the methods are well suited for use in pre-packaged kits for use by clinicians, researchers, and epidemiologists.

#### II. Diagnostic Kit

**[0045]** In another aspect, the invention provides a diagnostic kit for detecting the presence of a known or unknown adeno-associated virus (AAV) in a sample. Such a kit may contain a first set of 5' and 3' PCR primers specific for a signature region of the AAV nucleic acid sequence. Alternatively, or additionally, such a kit can contain a first set of 5' and 3' PCR primers specific for the 3.1 kb fragment which includes the full-length AAV capsid nucleic acid sequence identified herein (e.g., the AV1ns and AV2cas primers.) Optionally, a kit of the invention may further contain two or more additional sets of 5' and 3' primers, as described herein, and/or PCR probes. These primers and probes are used according to the present invention amplify signature regions of each AAV serotype, e.g., using quantitative PCR.

**[0046]** The invention further provides a kit useful for identifying an AAV serotype detected according to the method of the invention and/or for distinguishing novel AAV from known AAV. Such a kit may further include one or more restriction enzymes, standards for AAV serotypes providing their "signature restriction enzyme digestions analyses", and/or other means for determining the serotype of the AAV detected.

**[0047]** In addition, kits of the invention may include, instructions, a negative and/or positive control, containers, diluents and buffers for the sample, indicator charts for signature comparisons, disposable gloves, decontamination instructions, applicator sticks or containers, and sample preparator cups, as well as any desired reagents, including media, wash reagents and concentration reagents. Such reagents may be readily selected from among the reagents described herein, and from among conventional concentration reagents. In one desirable embodiment, the wash reagent is an isotonic saline solution which has been buffered to physiologic pH, such as phosphate buffered saline (PBS); the elution reagent is PBS containing 0.4 M NaCl, and the concentration reagents and devices. For example, one of skill in the art will recognize that reagents such as polyethylene glycol (PEG), or NH<sub>4</sub>SO<sub>4</sub> may be useful, or that devices such as filter devices. For example, a filter device with a 100 K membrane would concentrate rAAV.

**[0048]** The kits provided by the present invention are useful for performing the methods described herein, and for study of biodistribution, epidemiology, mode of transmission of novel AAV serotypes in human and NHPs.

**[0049]** Thus, the methods and kits of the invention permit detection, identification, and isolation of target viral sequences, particularly integrated viral sequences. The methods and kits are particularly well suited for use in detection, identification and isolation of AAV sequences, which may include novel AAV serotypes.

**[0050]** In one notable example, the method of the invention facilitated analysis of cloned AAV sequences by the inventors, which revealed heterogeneity of proviral sequences between cloned fragments from different animals, all of which were distinct from the known six AAV serotypes, with the majority of the variation localized to hypervariable regions of the capsid protein. Surprising divergence of AAV sequences was noted in clones isolated from single tissue sources, such as lymph node, from an individual rhesus monkey. This heterogeneity is best explained by apparent evolution of AAV sequence within individual animals due, in part, to extensive homologous recombination between a limited number of co-infecting parenteral viruses. These studies suggest sequence evolution of widely disseminated virus during the course of a natural AAV infection that presumably leads to the formation of swarms of quasispecies which differ from one another in the array of capsid hypervariable regions. This is the first example of rapid molecular evolution of a DNA virus in a way that formerly was thought to be restricted to RNA viruses.

**[0051]** Sequences of several novel AAV serotypes identified by the method of the invention and characterization of these serotypes is provided.

### III. Novel AAV Serotypes

#### A. Nucleic Acid Sequences

**[0052]** Nucleic acid sequences of novel AAV serotypes identified by the methods of the invention are provided. See, SEQ ID NO:1, 9 - 59, and 117 - 120, which are incorporated by reference herein. See also, Fig. 1 and the sequence listing.

**[0053]** For novel serotype AAV7, the full-length sequences, including the AAV 5' ITRs, capsid, rep, and AAV 3' ITRs are provided in SEQ ID NO:1.

**[0054]** For other novel AAV serotypes of the invention, the approximately 3.1 kb fragment isolated according to the method of the invention is provided. This fragment contains sequences encoding full-length capsid protein and all or part of the sequences encoding the rep protein. These sequences include the clones identified below.

**[0055]** For still other novel AAV serotypes, the signature region encoding the capsid protein is provided. For example, the AAV10 nucleic acid sequences of the invention include those illustrated in Fig. 1 [See, SEQ ID NO:117, which spans 255 bases]. The AAV11 nucleic acid sequences of the invention include the DNA sequences illustrated in Fig. 1 [See, SEQ ID NO:118 which spans 258 bases]. The AAV12 nucleic acid sequences of the invention include the DNA sequences illustrated in Fig. 1 [See, SEQ ID NO:119, which consists of 255 bases]. Using the methodology described above, further AAV10, AAV11 and AAV12 sequences can be readily identified and used for a variety of purposes, including those described for AAV7 and the other novel serotypes herein.

**[0056]** Figure 1 provides the non-human primate (NHP) AAV nucleic acid sequences of the invention in an alignment with the previously published AAV serotypes, AAV 1 [SEQ ID NO:6], AAV2 [SEQ ID NO:7], and AAV3 [SEQ ID NO:8]. These novel NHP sequences include those provided in the following Table I, which are identified by clone number:

Table 1

AAV Cap Sequence	Clone Number	Source		
		Species	Tissue	SEQ ID NO (DNA)
Rh.1	Clone 9 (AAV9)	Rhesus	Heart	5
Rh.2	Clone 43.1	Rhesus	MLN	39
Rh.3	Clone 43.5	Rhesus	MLN	40
Rh.4	Clone 43.12	Rhesus	MLN	41
Rh.5	Clone 43.20	Rhesus	MLN	42
Rh.6	Clone 43.21	Rhesus	MLN	43
Rh.7	Clone 43.23	Rhesus	MLN	44

Table 1 (cont'd)

5	Rh.8	Clone 43.25	Rhesus	MLN	45
	Rh.9	Clone 44.1	Rhesus	Liver	46
	Rh.10	Clone 44.2	Rhesus	Liver	59
10	Rh.11	Clone 44.5	Rhesus	Liver	47
	Rh.12	Clone 42.1B	Rhesus	MLN	30
	Rh.13	42.2	Rhesus	MLN	9
15	Rh.14	Clone 42.3A	Rhesus	MLN	32
	Rh.15	Clone 42.3B	Rhesus	MLN	36
	Rh.16	Clone 42.4	Rhesus	MLN	33
20	Rh.17	Clone 42.5A	Rhesus	MLN	34
	Rh.18	Clone 42.5B	Rhesus	MLN	29
25	Rh.19	Clone 42.6B	Rhesus	MLN	38
	Rh.20	Clone 42.8	Rhesus	MLN	27
	Rh.21	Clone 42.10	Rhesus	MLN	35
	Rh.22	Clone 42.11	Rhesus	MLN	37
30	Rh.23	Clone 42.12	Rhesus	MLN	58
	Rh.24	Clone 42.13	Rhesus	MLN	31
	Rh.25	Clone 42.15	Rhesus	MLN	28
	Rh.26	Clone 223.2	Rhesus	Liver	49
35	Rh.27	Clone 223.4	Rhesus	Liver	50
	Rh.28	Clone 223.5	Rhesus	Liver	51
	Rh.29	Clone 223.6	Rhesus	Liver	52
	Rh.30	Clone 223.7	Rhesus	Liver	53
40	Rh.31	Clone 223.10	Rhesus	Liver	48
	Rh.32	Clone C1	Rhesus	Spleen, Duo, Kid & Liver	19
	Rh.33	Clone C3	Rhesus		20
45	Rh.34	Clone C5	Rhesus		21
	Rh.35	Clone F1	Rhesus	Liver	22
	Rh.36	Clone F3	Rhesus		23
	Rh.37	Clone F5	Rhesus		24
50	Cy.1	Clone 1.3	Cyno	Blood	14
	Cy.2	Clone 13.3B	Cyno	Blood	15
	Cy.3	Clone 24.1	Cyno	Blood	16
	Cy.4	Clone 27.3	Cyno	Blood	17
55	Cy.5	Clone 7.2	Cyno	Blood	18
	Cy.6	Clone 16.3	Cyno	Blood	10

Table 1 (cont'd)

bb.1	Clone 29.3	Baboon	Blood	11
bb.2	Clone 29.5	Baboon	Blood	13
Ch.1	Clone A3.3	Chimp	Blood	57
Ch.2	Clone A3.4	Chimp	Blood	54
Ch.3	Clone A3.5	Chimp	Blood	55
Ch.4	Clone A3.7	Chimp	Blood	56

**[0057]** A novel NHP clone was made by splicing capsids fragments of two chimp adenoviruses into an AAV2 rep construct. This new clone, A3.1, is also termed Ch.5 [SEQ ID NO:20]. Additionally, the present invention includes two human AAV sequences, termed H6 [SEQ ID NO:25] and H2 [SEQ ID NO:26].

**[0058]** The AAV nucleic acid sequences of the invention further encompass the strand which is complementary to the strands provided in the sequences provided in Fig. 1 and the Sequence Listing [SEQ ID NO:1, 9 - 59, 117-120], nucleic acid sequences, as well as the RNA and cDNA sequences corresponding to the sequences provided in Fig. 1 and the Sequence Listing [SEQ ID NO:1, 9 - 59, 117-120], and their complementary strands. Also included in the nucleic acid sequences of the invention are natural variants and engineered modifications of the sequences of Fig. 1 and the Sequence Listing [SEQ ID NO:1, 9 - 59, 117-120], and their complementary strands. Such modifications include, for example, labels which are known in the art, methylation, and substitution of one or more of the naturally occurring nucleotides with a degenerate nucleotide.

**[0059]** Further included in this invention are nucleic acid sequences which are greater than 85%, preferably at least about 90%, more preferably at least about 95%, and most preferably at least about 98 to 99% identical or homologous to the sequences of the invention, including Fig. 1 and the Sequence Listing [SEQ ID NO:1, 9 - 59, 117-120]. These terms are as defined herein.

**[0060]** Also included within the invention are fragments of the novel AAV sequences identified by the method described herein. Suitable fragments are at least 15 nucleotides in length, and encompass functional fragments, i.e., fragments which are of biological interest. In one embodiment, these fragments are fragments of the novel sequences of Fig. 1 and the Sequence Listing [SEQ ID NO:1, 9 - 59, 117-120], their complementary strands, cDNA and RNA complementary thereto.

**[0061]** Examples of suitable fragments are provided with respect to the location of these fragments on AAV1, AAV2, or AAV7. However, using the alignment provided herein (obtained using the Clustal W program at default settings), or similar techniques for generating an alignment with other novel serotypes of the invention, one of skill in the art can readily identify the precise nucleotide start and stop codons for desired fragments.

**[0062]** Examples of suitable fragments include the sequences encoding the three variable proteins (vp) of the AAV capsid which are alternative splice variants: vp1 [e.g., nt 825 to 3049 of AAV7, SEQ ID NO: 1]; vp2 [e.g., nt 1234 - 3049 of AAV7, SEQ ID NO: 1]; and vp 3 [e.g., nt 1434 - 3049 of AAV7, SEQ ID NO:1]. It is notable that AAV7 has an unusual GTG start codon. With the exception of a few house-keeping genes, such a start codon has not previously been reported in DNA viruses. The start codons for vp1, vp2 and vp3 for other AAV serotypes have been believed to be such that they permit the cellular mechanism of the host cell in which they reside to produce vp1, vp2 and vp3 in a ratio of 10%:10%:80%, respectively, in order to permit efficient assembly of the virion. However, the AAV7 virion has been found to assemble efficiently even with this rare GTG start codon. Thus, the inventors anticipate this it is desirable to alter the start codon of the vp3 of other AAV serotypes to contain this rare GTG start codon, in order to improve packaging efficiency, to alter the virion structure and/or to alter location of epitopes (e.g., neutralizing antibody epitopes) of other AAV serotypes. The start codons may be altered using conventional techniques including, e.g., site directed mutagenesis. Thus, the present invention encompasses altered AAV virions of any selected serotype, composed of a vp 3, and/or optionally, vp 1 and/or vp2 having start codons altered to GTG.

**[0063]** Other suitable fragments of AAV, include a fragment containing the start codon for the AAV capsid protein [e.g., nt 468 to 3090 of AAV7, SEQ ID NO:1, nt 725 to 3090 of AAV7, SEQ ID NO: 1, and corresponding regions of the other AAV serotypes]. Still other fragments of AAV7 and the other novel AAV serotypes identified using the methods described herein include those encoding the rep proteins, including *rep 78* [e.g., initiation codon 334 of Fig. 1 for AAV7], *rep 68* [initiation codon nt 334 of Fig. 1 for AAV7], *rep 52* [initiation codon 1006 of Fig. 1 for AAV7], and *rep 40* [initiation codon 1006 of Fig. 1 for AAV7]. Other fragments of interest may include the AAV 5' inverted terminal repeats ITRs, [nt 1 to 107 of Fig. 1 for AAV7]; the AAV 3' ITRs [nt 4704 to 4721 of Fig. 1 for AAV7], P19 sequences, AAV P40 sequences, the rep binding site, and the terminal resolute site (TRS). Still other suitable fragments will be readily apparent to those of skill in the art. The corresponding regions in the other novel serotypes of the invention can be readily determined by reference to Figure 1, or by utilizing conventional alignment techniques with the sequences provided herein.

**[0064]** In addition to including the nucleic acid sequences provided in the figures and Sequence Listing, the present invention includes nucleic acid molecules and sequences which are designed to express the amino acid sequences, proteins and peptides of the AAV serotypes of the invention. Thus, the invention includes nucleic acid sequences which encode the following novel AAV amino acid sequences: C1 [SEQ ID NO:60], C2 [SEQ ID NO:61], C5 [SEQ ID NO:62], A3-3 [SEQ ID NO:66], A3-7 [SEQ ID NO:67], A3-4 [SEQ ID NO:68], A3-5 [SEQ ID NO: 69], 3.3b [SEQ ID NO: 62], 223.4 [SEQ ID NO: 73], 223-5 [SEQ ID NO:74], 223-10 [SEQ ID NO:75], 223-2 [SEQ ID NO:76], 223-7 [SEQ ID NO: 77], 223-6 [SEQ ID NO: 78], 44-1 [SEQ ID NO: 79], 44-5 [SEQ ID NO:80], 44-2 [SEQ ID NO:81], 42-15 [SEQ ID NO: 84], 42-8 [SEQ ID NO: 85], 42-13 [SEQ ID NO:86], 42-3A [SEQ ID NO:87], 42-4 [SEQ ID NO:88], 42-5A [SEQ ID NO: 89], 42-1B [SEQ ID NO:90], 42-5B [SEQ ID NO:91], 43-1 [SEQ ID NO: 92], 43-12 [SEQ ID NO: 93], 43-5 [SEQ ID NO: 94], 43-21 [SEQ ID NO:96], 43-25 [SEQ ID NO: 97], 43-20 [SEQ ID NO:99], 24.1 [SEQ ID NO: 101], 42.2 [SEQ ID NO: 102], 7.2 [SEQ ID NO: 103], 27.3 [SEQ ID NO: 104], 16.3 [SEQ ID NO: 105], 42.10 [SEQ ID NO: 106], 42-3B [SEQ ID NO: 107], 42-11 [SEQ ID NO: 108], F1 [SEQ ID NO: 109], F5 [SEQ ID NO: 110], F3 [SEQ ID NO:111], 42-6B [SEQ ID NO: 112], and/or 42-12 [SEQ ID NO: 113], and artificial AAV serotypes generated using these sequences and/or unique fragments thereof.

**[0065]** As used herein, artificial AAV serotypes include, without limitation, AAV with a non-naturally occurring capsid protein. Such an artificial capsid may be generated by any suitable technique, using a novel AAV sequence of the invention (e.g., a fragment of a vp1 capsid protein) in combination with heterologous sequences which may be obtained from another AAV serotype (known or novel), non-contiguous portions of the same AAV serotype, from a non-AAV viral source, or from a non-viral source. An artificial AAV serotype may be, without limitation, a chimeric AAV capsid, a recombinant AAV capsid, or a "humanized" AAV capsid.

#### B. AAV Amino Acid Sequences, Proteins and Peptides

**[0066]** The invention provides proteins and fragments thereof which are encoded by the nucleic acid sequences of the novel AAV serotypes identified herein, including, e.g., AAV7 [nt 825 to 3049 of AAV7, SEQ ID NO: 1] the other novel serotypes provided herein. Thus, the capsid proteins of the novel serotypes of the invention, including: H6 [SEQ ID NO: 25], H2 [SEQ ID NO: 26], 42-2 [SEQ ID NO:9], 42-8 [SEQ ID NO:27], 42-15 [SEQ ID NO:28], 42-5b [SEQ ID NO: 29], 42-1b [SEQ ID NO:30], 42-13 [SEQ ID NO: 31], 42-3a [SEQ ID NO: 32], 42-4 [SEQ ID NO:33], 42-5a [SEQ ID NO: 34], 42-10 [SEQ ID NO:35], 42-3b [SEQ ID NO: 36], 42-11 [SEQ ID NO: 37], 42-6b [SEQ ID NO:38], 43-1 [SEQ ID NO: 39], 43-5 [SEQ ID NO: 40], 43-12 [SEQ ID NO:41], 43-20 [SEQ ID NO:42], 43-21 [SEQ ID NO: 43], 43-23 [SEQ ID NO:44], 43-25 [SEQ ID NO: 45], 44.1 [SEQ ID NO:47], 44.5 [SEQ ID NO:47], 223.10 [SEQ ID NO:48], 223.2 [SEQ ID NO:49], 223.4 [SEQ ID NO:50], 223.5 [SEQ ID NO: 51], 223.6 [SEQ ID NO: 52], 223.7 [SEQ ID NO: 53], A3.4 [SEQ ID NO: 54], A3.5 [SEQ ID NO:55], A3.7 [SEQ ID NO: 56], A3.3 [SEQ ID NO:57], 42.12 [SEQ ID NO: 58], and 44.2 [SEQ ID NO: 59], can be readily generated using conventional techniques from the open reading frames provided for the above-listed clones.

**[0067]** The invention further encompasses AAV serotypes generated using sequences of the novel AAV serotypes of the invention, which are generated using synthetic, recombinant or other techniques known to those of skill in the art. The invention is not limited to novel AAV amino acid sequences, peptides and proteins expressed from the novel AAV nucleic acid sequences of the invention and encompasses amino acid sequences, peptides and proteins generated by other methods known in the art, including, e.g., by chemical synthesis, by other synthetic techniques, or by other methods. For example, the sequences of any of C1 [SEQ ID NO:60], C2 [SEQ ID NO:61], C5 [SEQ ID NO:62], A3-3 [SEQ ID NO:66], A3-7 [SEQ ID NO:67], A3-4 [SEQ ID NO:68], A3-5 [SEQ ID NO: 69], 3.3b [SEQ ID NO: 62], 223.4 [SEQ ID NO: 73], 223-5 [SEQ ID NO:74], 223-10 [SEQ ID NO:75], 223-2 [SEQ ID NO:76], 223-7 [SEQ ID NO: 77], 223-6 [SEQ ID NO: 78], 44-1 [SEQ ID NO: 79], 44-5 [SEQ ID NO:80], 44-2 [SEQ ID NO:81], 42-15 [SEQ ID NO: 84], 42-8 [SEQ ID NO: 85], 42-13 [SEQ ID NO:86], 42-3A [SEQ ID NO:87], 42-4 [SEQ ID NO:88], 42-5A [SEQ ID NO: 89], 42-1B [SEQ ID NO:90], 42-5B [SEQ ID NO:91], 43-1 [SEQ ID NO: 92], 43-12 [SEQ ID NO: 93], 43-5 [SEQ ID NO: 94], 43-21 [SEQ ID NO:96], 43-25 [SEQ ID NO: 97], 43-20 [SEQ ID NO:99], 24.1 [SEQ ID NO: 101], 42.2 [SEQ ID NO: 102], 7.2 [SEQ ID NO: 103], 27.3 [SEQ ID NO: 104], 16.3 [SEQ ID NO: 105], 42.10 [SEQ ID NO: 106], 42-3B [SEQ ID NO: 107], 42-11 [SEQ ID NO: 108], F1 [SEQ ID NO: 109], F5 [SEQ ID NO: 110], F3 [SEQ ID NO:111], 42-6B [SEQ ID NO: 112], and/or 42-12 [SEQ ID NO: 113] by be readily generated using a variety of techniques.

**[0068]** Suitable production techniques are well known to those of skill in the art. See, e.g., Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (Cold Spring Harbor, NY). Alternatively, peptides can also be synthesized by the well known solid phase peptide synthesis methods (Merrifield, *J Am. Chem. Soc.*, **85**:2149 (1962); Stewart and Young, *Solid Phase Peptide Synthesis* (Freeman, San Francisco, 1969) pp. 27-62). These and other suitable production methods are within the knowledge of those of skill in the art and are not a limitation of the present invention.

**[0069]** Particularly desirable proteins include the AAV capsid proteins, which are encoded by the nucleotide sequences identified above. The sequences of many of the capsid proteins of the invention are provided in an alignment in Fig.

2 and/or in the Sequence Listing, SEQ ID NO: 2 and 60 to 115, which is incorporated by reference herein. The AAV capsid is composed of three proteins, vp1, vp2 and vp3, which are alternative splice variants. The full-length sequence provided in these figures is that of vp1. Based on the numbering of the AAV7 capsid [SEQ ID NO:2], the sequences of vp2 span amino acid 138 - 737 of AAV7 and the sequences of vp3 span amino acids 203 - 737 of AAV7. With this information, one of skill in the art can readily determine the location of the vp2 and vp3 proteins for the other novel serotypes of the invention.

**[0070]** Other desirable proteins and fragments of the capsid protein include the constant and variable regions, located between hypervariable regions (HPV) and the sequences of the HPV regions themselves. An algorithm developed to determine areas of sequence divergence in AAV2 has yielded 12 hypervariable regions (HVR) of which 5 overlap or are part of the four previously described variable regions. [Chiorini *et al*, *J. Virol*, **73**:1309-19 (1999); Rutledge *et al*, *J. Virol.*, **72**:309-319] Using this algorithm and/or the alignment techniques described herein, the HVR of the novel AAV serotypes are determined. For example, with respect to the number of the AAV2 vp1 [SEQ ID NO:70], the HVR are located as follows: HVR1, aa 146-152; HVR2, aa 182-186; HVR3, aa 262-264; HVR4, aa 381-383; HVR5, aa 450-474; HVR6, aa 490-495; HVR7, aa500-504; HVR8, aa 514-522; HVR9, aa 534-555; HVR10, aa 581-594; HVR11, aa 658-667; and HVR12, aa 705-719. Utilizing an alignment prepared in accordance with conventional methods and the novel sequences provided herein [See, e.g., Figure 2], one can readily determine the location of the HVR in the novel AAV serotypes of the invention. For example, utilizing Figure 2, one can readily determine that for AAV7 [SEQ ID NO: 2], HVR1 is located at aa 146 - 152; HVR2 is located at 182-187; HVR3 is located at aa 263-266, HVR4 is located at aa 383-385, HVR5 is located at aa 451-475; HVR6 is located at aa 491-496 of AAV7; HVR7 is located at aa 501-505; HVR8 is located at aa 513-521; HVR9 is located at 533-554; HVR10 is located at aa 583-596; HVR11 is located at aa 660-669; HVR12 is located at aa 707-721. Using the information provided herein, the HVRs for the other novel serotypes of the invention can be readily determined.

**[0071]** In addition, within the capsid, amino acid cassettes of identity have been identified. These cassettes are of particular interest, as they are useful in constructing artificial serotypes, e.g., by replacing a HVR1 cassette of a selected serotype with an HVR1 cassette of another serotype. Certain of these cassettes of identity are noted in Fig. 2. See, Fig. 2, providing the Clustal X alignment, which has a ruler is displayed below the sequences, starting at 1 for the first residue position. The line above the ruler is used to mark strongly conserved positions. Three characters (\*, :, .) are used. "\*" indicates positions which have a single, fully conserved residue. ":" indicates that a "strong" group is fully conserved "." Indicates that a "weaker" group is fully conserved. These are all the positively scoring groups that occur in the Gonnet Pam250 matrix. The strong groups are defined as a strong score >0.5 and the weak groups are defined as weak score <0.5.

**[0072]** Additionally, examples of other suitable fragments of AAV capsids include, with respect to the numbering of AAV2 [SEQ ID NO:70], aa 24 - 42, aa 25 - 28; aa 81 - 85; aa133-165; aa 134 - 165; aa 137-143; aa 154-156; aa 194-208; aa 261-274; aa 262-274; aa 171-173; aa 413-417; aa 449-478; aa 494-525; aa 534-571; aa 581-601; aa 660-671; aa 709-723. Still other desirable fragments include, for example, in AAV7, amino acids 1 to 184 of SEQ ID NO:2, amino acids 199 to 259; amino acids 274 to 446; amino acids 603 to 659; amino acids 670 to 706; amino acids 724 to 736; aa 185 to 198; aa 260 to 273; aa447 to 477; aa495 to 602; aa660 to 669; and aa707 to 723. Still other desirable regions, based on the numbering of AAV7 [SEQ ID NO:2], are selected from among the group consisting of aa 185 to 198; aa 260 to 273; aa447 to 477; aa495 to 602; aa660 to 669; and aa707 to 723. Using the alignment provided herein performed using the Clustal X program at default settings, or using other commercially or publicly available alignment programs at default settings, one of skill in the art can readily determine corresponding fragments of the novel AAV capsids of the invention.

**[0073]** Other desirable proteins are the AAV rep proteins [aa 1 to 623 of SEQ ID NO:3 for AAV7] and functional fragments thereof, including, e.g., aa 1 to 171, aa 172 to 372, aa 373 to 444, aa 445 to 623 of SEQ ID NO:3, among others. Suitably, such fragments are at least 8 amino acids in length. See, Fig. 3. Comparable regions can be identified in the proteins of the other novel AAV of the invention, using the techniques described herein and those which are known in the art. In addition, fragments of other desired lengths may be readily utilized. Such fragments may be produced recombinantly or by other suitable means, e.g., chemical synthesis.

**[0074]** The sequences, proteins, and fragments of the invention may be produced by any suitable means, including recombinant production, chemical synthesis, or other synthetic means. Such production methods are within the knowledge of those of skill in the art and are not a limitation of the present invention.

#### IV. Production of rAAV with novel AAV capsids

**[0075]** The invention encompasses novel, wild-type AAV serotypes identified by the invention, the sequences of which wild-type AAV serotypes are free of DNA and/or cellular material with these viruses are associated in nature. In another aspect, the present invention provides molecules which utilize the novel AAV sequences of the invention, including fragments thereof, for production of molecules useful in delivery of a heterologous gene or other nucleic acid

sequences to a target cell.

**[0076]** The molecules of the invention which contain sequences of a novel AAV serotype of the invention include any genetic element (vector) which may be delivered to a host cell, e.g., naked DNA, a plasmid, phage, transposon, cosmid, episome, a protein in a non-viral delivery vehicle (e.g., a lipid-based carrier), virus, etc. which transfer the sequences carried thereon. The selected vector may be delivered by any suitable method, including transfection, electroporation, liposome delivery, membrane fusion techniques, high velocity DNA-coated pellets, viral infection and protoplast fusion. The methods used to construct any embodiment of this invention are known to those with skill in nucleic acid manipulation and include genetic engineering, recombinant engineering, and synthetic techniques. See, e.g., Sambrook et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY.

**[0077]** In one embodiment, the vectors of the invention contain sequences encoding a novel AAV capsid of the invention (e.g., AAV7 capsid, AAV 44-2 (rh.10), an AAV10 capsid, an AAV11 capsid, an AAV12 capsid), or a fragment of one or more of these AAV capsids. Alternatively, the vectors may contain the capsid protein, or a fragment thereof, itself.

**[0078]** Optionally, vectors of the invention may contain sequences encoding AAV rep proteins. Such *rep* sequences may be from the same AAV serotype which is providing the *cap* sequences. Alternatively, the present invention provides vectors in which the *rep* sequences are from an AAV serotype which differs from that which is providing the *cap* sequences. In one embodiment, the *rep* and *cap* sequences are expressed from separate sources (e.g., separate vectors, or a host cell and a vector). In another embodiment, these *rep* sequences are expressed from the same source as the *cap* sequences. In this embodiment, the *rep* sequences may be fused in frame to *cap* sequences of a different AAV serotype to form a chimeric AAV vector. Optionally, the vectors of the invention further contain a minigene comprising a selected transgene which is flanked by AAV 5' ITR and AAV 3' ITR.

**[0079]** Thus, in one embodiment, the vectors described herein contain nucleic acid sequences encoding an intact AAV capsid which may be from a single AAV serotype (e.g., AAV7 or another novel AAV). Alternatively, these vectors contain sequences encoding artificial capsids which contain one or more fragments of the AAV7 (or another novel AAV) capsid fused to heterologous AAV or non-AAV capsid proteins (or fragments thereof). These artificial capsid proteins are selected from non-contiguous portions of the AAV7 (or another novel AAV) capsid or from capsids of other AAV serotypes. For example, it may be desirable to modify the coding regions of one or more of the AAV vp1, e.g., in one or more of the hypervariable regions (i.e., HPV1-12), or vp2, and/or vp3. In another example, it may be desirable to alter the start codon of the vp3 protein to GTG. These modifications may be to increase expression, yield, and/or to improve purification in the selected expression systems, or for another desired purpose (e.g., to change tropism or alter neutralizing antibody epitopes).

**[0080]** The vectors described herein, e.g., a plasmid, are useful for a variety of purposes, but are particularly well suited for use in production of a rAAV containing a capsid comprising AAV sequences or a fragment thereof. These vectors, including rAAV, their elements, construction, and uses are described in detail herein.

**[0081]** In one aspect, the invention provides a method of generating a recombinant adeno-associated virus (AAV) having an AAV serotype 7 (or another novel AAV) capsid, or a portion thereof. Such a method involves culturing a host cell which contains a nucleic acid sequence encoding an adeno-associated virus (AAV) serotype 7 (or another novel AAV) capsid protein, or fragment thereof, as defined herein; a functional *rep* gene; a minigene composed of, at a minimum, AAV inverted terminal repeats (ITRs) and a transgene; and sufficient helper functions to permit packaging of the minigene into the AAV7 (or another novel AAV) capsid protein.

**[0082]** The components required to be cultured in the host cell to package an AAV minigene in an AAV capsid may be provided to the host cell in *trans*. Alternatively, any one or more of the required components (e.g., minigene, *rep* sequences, *cap* sequences, and/or helper functions) may be provided by a stable host cell which has been engineered to contain one or more of the required components using methods known to those of skill in the art. Most suitably, such a stable host cell will contain the required component(s) under the control of an inducible promoter. However, the required component(s) may be under the control of a constitutive promoter. Examples of suitable inducible and constitutive promoters are provided herein, in the discussion of regulatory elements suitable for use with the transgene. In still another alternative, a selected stable host cell may contain selected component(s) under the control of a constitutive promoter and other selected component(s) under the control of one or more inducible promoters. For example, a stable host cell may be generated which is derived from 293 cells (which contain E1 helper functions under the control of a constitutive promoter), but which contains the *rep* and/or *cap* proteins under the control of inducible promoters. Still other stable host cells may be generated by one of skill in the art.

**[0083]** The minigene, *rep* sequences, *cap* sequences, and helper functions required for producing the rAAV of the invention may be delivered to the packaging host cell in the form of any genetic element which transfer the sequences carried thereon. The selected genetic element may be delivered by any suitable method, including those described herein. The methods used to construct any embodiment of this invention are known to those with skill in nucleic acid manipulation and include genetic engineering, recombinant engineering, and synthetic techniques. See, e.g., Sambrook et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY. Similarly,

methods of generating rAAV virions are well known and the selection of a suitable method is not a limitation on the present invention. See, e.g., K. Fisher et al, *J. Virol.*, **70**:520-532 (1993) and US Patent 5,478,745.

## A. The Minigene

**[0084]** The minigene is composed of, at a minimum, a transgene and its regulatory sequences, and 5' and 3' AAV inverted terminal repeats (ITRs). It is this minigene which is packaged into a capsid protein and delivered to a selected host cell.

### 1. The transgene

**[0085]** The transgene is a nucleic acid sequence, heterologous to the vector sequences flanking the transgene, which encodes a polypeptide, protein, or other product, of interest. The nucleic acid coding sequence is operatively linked to regulatory components in a manner which permits transgene transcription, translation, and/or expression in a host cell.

**[0086]** The composition of the transgene sequence will depend upon the use to which the resulting vector will be put. For example, one type of transgene sequence includes a reporter sequence, which upon expression produces a detectable signal. Such reporter sequences include, without limitation, DNA sequences encoding  $\beta$ -lactamase,  $\beta$ -galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyl-transferase (CAT), luciferase, membrane bound proteins including, for example, CD2, CD4, CD8, the influenza hemagglutinin protein, and others well known in the art, to which high affinity antibodies directed thereto exist or can be produced by conventional means, and fusion proteins comprising a membrane bound protein appropriately fused to an antigen tag domain from, among others, hemagglutinin or Myc.

**[0087]** These coding sequences, when associated with regulatory elements which drive their expression, provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence or other spectrographic assays, fluorescent activating cell sorting assays and immunological assays, including enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and immunohistochemistry. For example, where the marker sequence is the LacZ gene, the presence of the vector carrying the signal is detected by assays for beta-galactosidase activity. Where the transgene is green fluorescent protein or luciferase, the vector carrying the signal may be measured visually by color or light production in a luminometer.

**[0088]** However, desirably, the transgene is a non-marker sequence encoding a product which is useful in biology and medicine, such as proteins, peptides, RNA, enzymes, or catalytic RNAs. Desirable RNA molecules include tRNA, dsRNA, ribosomal RNA, catalytic RNAs, and antisense RNAs. One example of a useful RNA sequence is a sequence which extinguishes expression of a targeted nucleic acid sequence in the treated animal.

**[0089]** The transgene may be used to correct or ameliorate gene deficiencies, which may include deficiencies in which normal genes are expressed at less than normal levels or deficiencies in which the functional gene product is not expressed. A preferred type of transgene sequence encodes a therapeutic protein or polypeptide which is expressed in a host cell. The invention further includes using multiple transgenes, e.g., to correct or ameliorate a gene defect caused by a multi-subunit protein. In certain situations, a different transgene may be used to encode each subunit of a protein, or to encode different peptides or proteins. This is desirable when the size of the DNA encoding the protein subunit is large, e.g., for an immunoglobulin, the platelet-derived growth factor, or a dystrophin protein. In order for the cell to produce the multi-subunit protein, a cell is infected with the recombinant virus containing each of the different subunits. Alternatively, different subunits of a protein may be encoded by the same transgene. In this case, a single transgene includes the DNA encoding each of the subunits, with the DNA for each subunit separated by an internal ribozyme entry site (IRES). This is desirable when the size of the DNA encoding each of the subunits is small, e.g., the total size of the DNA encoding the subunits and the IRES is less than five kilobases. As an alternative to an IRES, the DNA may be separated by sequences encoding a 2A peptide, which self-cleaves in a post-translational event. See, e.g., M.L. Donnelly, *et al*, *J. Gen. Virol.*, **78**(Pt 1):13-21 (Jan 1997); Furler, S., *et al*, *Gene Ther.*, **8**(11):864-873 (June 2001); Klump H., *et al.*, *Gene Ther.*, **8**(10):811-817 (May 2001). This 2A peptide is significantly smaller than an IRES, making it well suited for use when space is a limiting factor. However, the selected transgene may encode any biologically active product or other product, e.g., a product desirable for study.

**[0090]** Suitable transgenes may be readily selected by one of skill in the art. The selection of the transgene is not considered to be a limitation of this invention.

### 2. Regulatory Elements

**[0091]** In addition to the major elements identified above for the minigene, the vector also includes conventional control elements necessary which are operably linked to the transgene in a manner which permits its transcription,

translation and/or expression in a cell transfected with the plasmid vector or infected with the virus produced by the invention. As used herein, "operably linked" sequences include both expression control sequences that are contiguous with the gene of interest and expression control sequences that act in *trans* or at a distance to control the gene of interest.

**[0092]** Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation (polyA) signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (i.e., Kozak consensus sequence); sequences that enhance protein stability; and when desired, sequences that enhance secretion of the encoded product. A great number of expression control sequences, including promoters which are native, constitutive, inducible and/or tissue-specific, are known in the art and may be utilized.

**[0093]** Examples of constitutive promoters include, without limitation, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) [see, e.g., Boshart *et al*, *Cell*, **41**:521-530 (1985)], the SV40 promoter, the dihydrofolate reductase promoter, the  $\beta$ -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1 $\alpha$  promoter [Invitrogen].

**[0094]** Inducible promoters allow regulation of gene expression and can be regulated by exogenously supplied compounds, environmental factors such as temperature, or the presence of a specific physiological state, e.g., acute phase, a particular differentiation state of the cell, or in replicating cells only. Inducible promoters and inducible systems are available from a variety of commercial sources, including, without limitation, Invitrogen, Clontech and Ariad. Many other systems have been described and can be readily selected by one of skill in the art. Examples of inducible promoters regulated by exogenously supplied promoters include the zinc-inducible sheep metallothioneine (MT) promoter, the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter, the T7 polymerase promoter system [WO 98/10088]; the ecdysone insect promoter [No *et al*, *Proc. Natl. Acad. Sci. USA*, **93**:3346-3351 (1996)], the tetracycline-repressible system [Gossen *et al*, *Proc. Natl. Acad. Sci. USA*, **89**:5547-5551 (1992)], the tetracycline-inducible system [Gossen *et al*, *Science*, **268**:1766-1769 (1995), see also Harvey *et al*, *Curr. Opin. Chem. Biol.*, **2**:512-518 (1998)], the RU486-inducible system [Wang *et al*, *Nat. Biotech.*, **15**:239-243 (1997) and Wang *et al*, *Gene Ther.*, **4**:432-441 (1997)] and the rapamycin-inducible system [Magari *et al*, *J. Clin. Invest.*, **100**:2865-2872 (1997)]. Still other types of inducible promoters which may be useful in this context are those which are regulated by a specific physiological state, e.g., temperature, acute phase, a particular differentiation state of the cell, or in replicating cells only.

**[0095]** In another embodiment, the native promoter for the transgene will be used. The native promoter may be preferred when it is desired that expression of the transgene should mimic the native expression. The native promoter may be used when expression of the transgene must be regulated temporally or developmentally, or in a tissue-specific manner, or in response to specific transcriptional stimuli. In a further embodiment, other native expression control elements, such as enhancer elements, polyadenylation sites or Kozak consensus sequences may also be used to mimic the native expression.

**[0096]** Another embodiment of the transgene includes a transgene operably linked to a tissue-specific promoter. For instance, if expression in skeletal muscle is desired, a promoter active in muscle should be used. These include the promoters from genes encoding skeletal  $\beta$ -actin, myosin light chain 2A, dystrophin, muscle creatine kinase, as well as synthetic muscle promoters with activities higher than naturally-occurring promoters (see Li *et al*, *Nat. Biotech.*, **17**:241-245 (1999)). Examples of promoters that are tissue-specific are known for liver (albumin, Miyatake *et al*, *J. Virol.*, **71**:5124-32 (1997); hepatitis B virus core promoter, Sandig *et al*, *Gene Ther.*, **3**:1002-9 (1996); alpha-fetoprotein (AFP), Arbutnot *et al*, *Hum. Gene Ther.*, **7**:1503-14 (1996)), bone osteocalcin (Stein *et al*, *Mol. Biol. Rep.*, **24**:185-96 (1997)); bone sialoprotein (Chen *et al*, *J. Bone Miner. Res.*, **11**:654-64 (1996)), lymphocytes (CD2, Hansal *et al*, *J. Immunol.*, **161**:1063-8 (1998); immunoglobulin heavy chain; T cell receptor  $\alpha$  chain), neuronal such as neuron-specific enolase (NSE) promoter (Andersen *et al*, *Cell. Mol. Neurobiol.*, **13**:503-15 (1993)), neurofilament light-chain gene (Piccioli *et al*, *Proc. Natl. Acad. Sci. USA*, **88**:5611-5 (1991)), and the neuron-specific vgf gene (Piccioli *et al*, *Neuron*, **15**:373-84 (1995)), among others.

**[0097]** Optionally, plasmids carrying therapeutically useful transgenes may also include selectable markers or reporter genes may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. Such selectable reporters or marker genes (preferably located outside the viral genome to be rescued by the method of the invention) can be used to signal the presence of the plasmids in bacterial cells, such as ampicillin resistance. Other components of the plasmid may include an origin of replication. Selection of these and other promoters and vector elements are conventional and many such sequences are available [see, e.g., Sambrook *et al*, and references cited therein].

**[0098]** The combination of the transgene, promoter/enhancer, and 5' and 3' ITRs is referred to as a "minigene" for ease of reference herein. Provided with the teachings of this invention, the design of such a minigene can be made by resort to conventional techniques.

## 3. Delivery of the Minigene to a Packaging Host Cell

**[0099]** The minigene can be carried on any suitable vector, e.g., a plasmid, which is delivered to a host cell. The plasmids useful in this invention may be engineered such that they are suitable for replication and, optionally, integration in prokaryotic cells, mammalian cells, or both. These plasmids (or other vectors carrying the 5' AAV ITR-heterologous molecule-3' ITR) contain sequences permitting replication of the minigene in eukaryotes and/or prokaryotes and selection markers for these systems. Selectable markers or reporter genes may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. The plasmids may also contain certain selectable reporters or marker genes that can be used to signal the presence of the vector in bacterial cells, such as ampicillin resistance. Other components of the plasmid may include an origin of replication and an amplicon, such as the amplicon system employing the Epstein Barr virus nuclear antigen. This amplicon system, or other similar amplicon components permit high copy episomal replication in the cells. Preferably, the molecule carrying the minigene is transfected into the cell, where it may exist transiently. Alternatively, the minigene (carrying the 5' AAV ITR-heterologous molecule-3' ITR) may be stably integrated into the genome of the host cell, either chromosomally or as an episome. In certain embodiments, the minigene may be present in multiple copies, optionally in head-to-head, head-to-tail, or tail-to-tail concatamers. Suitable transfection techniques are known and may readily be utilized to deliver the minigene to the host cell.

**[0100]** Generally, when delivering the vector comprising the minigene by transfection, the vector is delivered in an amount from about 5 µg to about 100 µg DNA, and preferably about 10 to about 50 µg DNA to about  $1 \times 10^4$  cells to about  $1 \times 10^{13}$  cells, and preferably about  $10^5$  cells. However, the relative amounts of vector DNA to host cells may be adjusted, taking into consideration such factors as the selected vector, the delivery method and the host cells selected.

B. *Rep* and *Cap* Sequences

**[0101]** In addition to the minigene, the host cell contains the sequences which drive expression of the novel AAV capsid protein (e.g., AAV7 or other novel AAV capsid or an artificial capsid protein comprising a fragment of one or more of these capsids) in the host cell and *rep* sequences of the same serotype as the serotype of the AAV ITRs found in the minigene. The AAV *cap* and *rep* sequences may be independently obtained from an AAV source as described above and may be introduced into the host cell in any manner known to one in the art as described above. Additionally, when pseudotyping a novel AAV capsid of the invention, the sequences encoding each of the essential *rep* proteins may be supplied by the same AAV serotype, or the sequences encoding the *rep* proteins may be supplied by different AAV serotypes (e.g., AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, or one of the novel serotypes identified herein). For example, the *rep78/68* sequences may be from AAV2, whereas the *rep52/40* sequences may from AAV1.

**[0102]** In one embodiment, the host cell stably contains the capsid protein under the control of a suitable promoter, such as those described above. Most desirably, in this embodiment, the capsid protein is expressed under the control of an inducible promoter. In another embodiment, the capsid protein is supplied to the host cell in *trans*. When delivered to the host cell in *trans*, the capsid protein may be delivered via a plasmid which contains the sequences necessary to direct expression of the selected capsid protein in the host cell. Most desirably, when delivered to the host cell in *trans*, the plasmid carrying the capsid protein also carries other sequences required for packaging the rAAV, e.g., the *rep* sequences.

**[0103]** In another embodiment, the host cell stably contains the *rep* sequences under the control of a suitable promoter, such as those described above. Most desirably, in this embodiment, the essential *rep* proteins are expressed under the control of an inducible promoter. In another embodiment, the *rep* proteins are supplied to the host cell in *trans*. When delivered to the host cell in *trans*, the *rep* proteins may be delivered via a plasmid which contains the sequences necessary to direct expression of the selected *rep* proteins in the host cell. Most desirably, when delivered to the host cell in *trans*, the plasmid carrying the capsid protein also carries other sequences required for packaging the rAAV, e.g., the *rep* and *cap* sequences.

**[0104]** Thus, in one embodiment, the *rep* and *cap* sequences may be transfected into the host cell on a single nucleic acid molecule and exist stably in the cell as an episome. In another embodiment, the *rep* and *cap* sequences are stably integrated into the genome of the cell. Another embodiment has the *rep* and *cap* sequences transiently expressed in the host cell. For example, a useful nucleic acid molecule for such transfection comprises, from 5' to 3', a promoter, an optional spacer interposed between the promoter and the start site of the *rep* gene sequence, an AAV *rep* gene sequence, and an AAV *cap* gene sequence.

**[0105]** Optionally, the *rep* and/or *cap* sequences may be supplied on a vector that contains other DNA sequences that are to be introduced into the host cells. For instance, the vector may contain the rAAV construct comprising the minigene. The vector may comprise one or more of the genes encoding the helper functions, e.g., the adenoviral proteins E1, E2a, and E4ORF6, and the gene for VAI RNA.

**[0106]** Preferably, the promoter used in this construct may be any of the constitutive, inducible or native promoters

known to one of skill in the art or as discussed above. In one embodiment, an AAV P5 promoter sequence is employed. The selection of the AAV to provide any of these sequences does not limit the invention.

**[0107]** In another preferred embodiment, the promoter for *rep* is an inducible promoter, as are discussed above in connection with the transgene regulatory elements. One preferred promoter for *rep* expression is the T7 promoter. The vector comprising the *rep* gene regulated by the T7 promoter and the *cap* gene, is transfected or transformed into a cell which either constitutively or inducibly expresses the T7 polymerase. See WO 98/10088, published March 12, 1998.

**[0108]** The spacer is an optional element in the design of the vector. The spacer is a DNA sequence interposed between the promoter and the *rep* gene ATG start site. The spacer may have any desired design; that is, it may be a random sequence of nucleotides, or alternatively, it may encode a gene product, such as a marker gene. The spacer may contain genes which typically incorporate start/stop and polyA sites. The spacer may be a non-coding DNA sequence from a prokaryote or eukaryote, a repetitive non-coding sequence, a coding sequence without transcriptional controls or a coding sequence with transcriptional controls. Two exemplary sources of spacer sequences are the  $\lambda$  phage ladder sequences or yeast ladder sequences, which are available commercially, e.g., from Gibco or Invitrogen, among others. The spacer may be of any size sufficient to reduce expression of the *rep78* and *rep68* gene products, leaving the *rep52*, *rep40* and *cap* gene products expressed at normal levels. The length of the spacer may therefore range from about 10 bp to about 10.0 kbp, preferably in the range of about 100 bp to about 8.0 kbp. To reduce the possibility of recombination, the spacer is preferably less than 2 kbp in length; however, the invention is not so limited.

**[0109]** Although the molecule(s) providing *rep* and *cap* may exist in the host cell transiently (i.e., through transfection), it is preferred that one or both of the *rep* and *cap* proteins and the promoter(s) controlling their expression be stably expressed in the host cell, e.g., as an episome or by integration into the chromosome of the host cell. The methods employed for constructing embodiments of this invention are conventional genetic engineering or recombinant engineering techniques such as those described in the references above. While this specification provides illustrative examples of specific constructs, using the information provided herein, one of skill in the art may select and design other suitable constructs, using a choice of spacers, P5 promoters, and other elements, including at least one translational start and stop signal, and the optional addition of polyadenylation sites.

**[0110]** In another embodiment of this invention, the *rep* or *cap* protein may be provided stably by a host cell.

#### C. The Helper Functions

**[0111]** The packaging host cell also requires helper functions in order to package the rAAV of the invention. Optionally, these functions may be supplied by a herpesvirus. Most desirably, the necessary helper functions are each provided from a human or non-human primate adenovirus source, such as those described above and/or are available from a variety of sources, including the American Type Culture Collection (ATCC), Manassas, VA (US). In one currently preferred embodiment, the host cell is provided with and/or contains an E1a gene product, an E1b gene product, an E2a gene product, and/or an E4 ORF6 gene product. The host cell may contain other adenoviral genes such as VAI RNA, but these genes are not required. In a preferred embodiment, no other adenovirus genes or gene functions are present in the host cell.

**[0112]** By "adenoviral DNA which expresses the E1a gene product", it is meant any adenovirus sequence encoding E1a or any functional E1a portion. Adenoviral DNA which expresses the E2a gene product and adenoviral DNA which expresses the E4 ORF6 gene products are defined similarly. Also included are any alleles or other modifications of the adenoviral gene or functional portion thereof. Such modifications may be deliberately introduced by resort to conventional genetic engineering or mutagenic techniques to enhance the adenoviral function in some manner, as well as naturally occurring allelic variants thereof. Such modifications and methods for manipulating DNA to achieve these adenovirus gene functions are known to those of skill in the art.

**[0113]** The adenovirus E1a, E1b, E2a, and/or E4ORF6 gene products, as well as any other desired helper functions, can be provided using any means that allows their expression in a cell. Each of the sequences encoding these products may be on a separate vector, or one or more genes may be on the same vector. The vector may be any vector known in the art or disclosed above, including plasmids, cosmids and viruses. Introduction into the host cell of the vector may be achieved by any means known in the art or as disclosed above, including transfection, infection, electroporation, liposome delivery, membrane fusion techniques, high velocity DNA-coated pellets, viral infection and protoplast fusion, among others. One or more of the adenoviral genes may be stably integrated into the genome of the host cell, stably expressed as episomes, or expressed transiently. The gene products may all be expressed transiently, on an episome or stably integrated, or some of the gene products may be expressed stably while others are expressed transiently. Furthermore, the promoters for each of the adenoviral genes may be selected independently from a constitutive promoter, an inducible promoter or a native adenoviral promoter. The promoters may be regulated by a specific physiological state of the organism or cell (i.e., by the differentiation state or in replicating or quiescent cells) or by exogenously-added factors, for example.

## D. Host Cells And Packaging Cell Lines

**[0114]** The host cell itself may be selected from any biological organism, including prokaryotic (e.g., bacterial) cells, and eukaryotic cells, including, insect cells, yeast cells and mammalian cells. Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549, WEHI, 3T3, 10T1/2, BHK, MDCK, COS 1, COS 7, BSC 1, BSC 40, BMT 10, VERO, WI38, HeLa, 293 cells (which express functional adenoviral E1), Saos, C2C12, L cells, HT1080, HepG2 and primary fibroblast, hepatocyte and myoblast cells derived from mammals including human, monkey, mouse, rat, rabbit, and hamster. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, i.e., fibroblast, hepatocyte, tumor cell, etc. The most desirable cells do not carry any adenovirus gene other than E1, E2a and/or E4 ORF6; nor do they contain any other virus gene which could result in homologous recombination of a contaminating virus during the production of rAAV; and it is capable of infection or transfection of DNA and expression of the transfected DNA. In a preferred embodiment, the host cell is one that has *rep* and *cap* stably transfected in the cell.

**[0115]** One host cell useful in the present invention is a host cell stably transformed with the sequences encoding *rep* and *cap*, and which is transfected with the adenovirus E1, E2a, and E4ORF6 DNA and a construct carrying the minigene as described above. Stable *rep* and/or *cap* expressing cell lines, such as B-50 (PCT/US98/19463), or those described in U.S. Patent No. 5,658,785, may also be similarly employed. Another desirable host cell contains the minimum adenoviral DNA which is sufficient to express E4 ORF6. Yet other cell lines can be constructed using the novel AAV *rep* and/or novel AAV *cap* sequences of the invention.

**[0116]** The preparation of a host cell according to this invention involves techniques such as assembly of selected DNA sequences. This assembly may be accomplished utilizing conventional techniques. Such techniques include cDNA and genomic cloning, which are well known and are described in Sambrook et al., cited above, use of overlapping oligonucleotide sequences of the adenovirus and AAV genomes, combined with polymerase chain reaction, synthetic methods, and any other suitable methods which provide the desired nucleotide sequence.

**[0117]** Introduction of the molecules (as plasmids or viruses) into the host cell may also be accomplished using techniques known to the skilled artisan and as discussed throughout the specification. In preferred embodiment, standard transfection techniques are used, e.g.,  $\text{CaPO}_4$  transfection or electroporation, and/or infection by hybrid adenovirus/AAV vectors into cell lines such as the human embryonic kidney cell line HEK 293 (a human kidney cell line containing functional adenovirus E1 genes which provides *trans*-acting E1 proteins).

**[0118]** These novel AAV-based vectors which are generated by one of skill in the art are beneficial for gene delivery to selected host cells and gene therapy patients since no neutralization antibodies to AAV7 have been found in the human population. Further, early studies show no neutralizing antibodies in cyno monkey and chimpanzee populations, and less than 15% cross-reactivity of AAV 7 in rhesus monkeys, the species from which the serotype was isolated. One of skill in the art may readily prepare other rAAV viral vectors containing the AAV7 capsid proteins provided herein using a variety of techniques known to those of skill in the art. One may similarly prepare still other rAAV viral vectors containing AAV7 sequence and AAV capsids of another serotype. Similar advantages are conferred by the vectors based on the other novel AAV of the invention.

**[0119]** Thus, one of skill in the art will readily understand that the AAV7 sequences of the invention can be readily adapted for use in these and other viral vector systems for *in vitro*, *ex vivo* or *in vivo* gene delivery. Similarly, one of skill in the art can readily select other fragments of the novel AAV genome of the invention for use in a variety of rAAV and non-rAAV vector systems. Such vectors systems may include, e.g., lentiviruses, retroviruses, poxviruses, vaccinia viruses, and adenoviral systems, among others. Selection of these vector systems is not a limitation of the present invention.

**[0120]** Thus, the invention further provides vectors generated using the nucleic acid and amino acid sequences of the novel AAV of the invention. Such vectors are useful for a variety of purposes, including for delivery of therapeutic molecules and for use in vaccine regimens. Particularly desirable for delivery of therapeutic molecules are recombinant AAV containing capsids of the novel AAV of the invention. These, or other vector constructs containing novel AAV sequences of the invention may be used in vaccine regimens, e.g., for co-delivery of a cytokine, or for delivery of the immunogen itself.

## V. Recombinant Viruses And Uses Thereof

**[0121]** Using the techniques described herein, one of skill in the art may generate a rAAV having a capsid of a novel serotype of the invention, or a novel capsid containing one or more novel fragments of an AAV serotype identified by the method of the invention. In one embodiment, a full-length capsid from a single serotype, e.g., AAV7 [SEQ ID NO: 2] can be utilized. In another embodiment, a full-length capsid may be generated which contains one or more fragments of a novel serotype of the invention fused in frame with sequences from another selected AAV serotype. For example, a rAAV may contain one or more of the novel hypervariable region sequences of an AAV serotype of the invention.

Alternatively, the unique AAV serotypes of the invention may be used in constructs containing other viral or non-viral sequences.

**[0122]** It will be readily apparent to one of skill in the art one embodiment, that certain serotypes of the invention will be particularly well suited for certain uses. For example, vectors based on AAV7 capsids of the invention are particularly well suited for use in muscle; whereas vectors based on rh.10 (44-2) capsids of the invention are particularly well suited for use in lung. Uses of such vectors are not so limited and one of skill in the art may utilize these vectors for delivery to other cell types, tissues or organs. Further, vectors based upon other capsids of the invention may be used for delivery to these or other cells, tissues or organs.

#### A. Delivery of Transgene

**[0123]** In another aspect, the present invention provides a method for delivery of a transgene to a host which involves transfecting or infecting a selected host cell with a vector generated with the sequences of the AAV of the invention. Methods for delivery are well known to those of skill in the art and are not a limitation of the present invention.

**[0124]** In one desirable embodiment, the invention provides a method for AAV-mediated delivery of a transgene to a host. This method involves transfecting or infecting a selected host cell with a recombinant viral vector containing a selected transgene under the control of sequences which direct expression thereof and AAV capsid proteins.

**[0125]** Optionally, a sample from the host may be first assayed for the presence of antibodies to a selected AAV serotype. A variety of assay formats for detecting neutralizing antibodies are well known to those of skill in the art. The selection of such an assay is not a limitation of the present invention. See, e.g., Fisher et al, *Nature Med.*, **3**(3):306-312 (March 1997) and W. C. Manning et al, *Human Gene Therapy*, **9**:477-485 (March 1, 1998). The results of this assay may be used to determine which AAV vector containing capsid proteins of a particular serotype are preferred for delivery, e.g., by the absence of neutralizing antibodies specific for that capsid serotype.

**[0126]** In one aspect of this method, the delivery of vector with a selected AAV capsid proteins may precede or follow delivery of a gene via a vector with a different serotype AAV capsid protein. Similarly, the delivery of vector with other novel AAV capsid proteins of the invention may precede or follow delivery of a gene via a vector with a different serotype AAV capsid protein. Thus, gene delivery via rAAV vectors may be used for repeat gene delivery to a selected host cell. Desirably, subsequently administered rAAV vectors carry the same transgene as the first rAAV vector, but the subsequently administered vectors contain capsid proteins of serotypes which differ from the first vector. For example, if a first vector has AAV7 capsid proteins [SEQ ID NO:2], subsequently administered vectors may have capsid proteins selected from among the other serotypes, including AAV1, AAV2, AAV3A, AAV3B, AAV4, AAV6, AAV10, AAV11, and AAV12, or any of the other novel AAV capsids identified herein including, without limitation: A3.1, H2, H6, C1, C2, C5, A3-3, A3-7, A3-4, A3-5, 3.3b, 223.4, 223-5, 223-10, 223-2, 223-7, 223-6, 44-1, 44-5, 44-2, 42-15, 42-8, 42-13, 42-3A, 42-4, 42-5A, 42-1B, 42-5B, 43-1, 43-12, 43-5, 43-21, 43-25, 43-20, 24.1, 42.2, 7.2, 27.3, 16.3, 42.10, 42-3B, 42-11, F1, F5, F3, 42-6B, and/or 42-12.

**[0127]** The above-described recombinant vectors may be delivered to host cells according to published methods. The rAAV, preferably suspended in a physiologically compatible carrier, may be administered to a human or non-human mammalian patient. Suitable carriers may be readily selected by one of skill in the art in view of the indication for which the transfer virus is directed. For example, one suitable carrier includes saline, which may be formulated with a variety of buffering solutions (e.g., phosphate buffered saline). Other exemplary carriers include sterile saline, lactose, sucrose, calcium phosphate, gelatin, dextran, agar, pectin, peanut oil, sesame oil, and water. The selection of the carrier is not a limitation of the present invention.

**[0128]** Optionally, the compositions of the invention may contain, in addition to the rAAV and carrier(s), other conventional pharmaceutical ingredients, such as preservatives, or chemical stabilizers. Suitable exemplary preservatives include chlorobutanol, potassium sorbate, sorbic acid, sulfur dioxide, propyl gallate, the parabens, ethyl vanillin, glycerin, phenol, and parachlorophenol. Suitable chemical stabilizers include gelatin and albumin.

**[0129]** The viral vectors are administered in sufficient amounts to transfect the cells and to provide sufficient levels of gene transfer and expression to provide a therapeutic benefit without undue adverse effects, or with medically acceptable physiological effects, which can be determined by those skilled in the medical arts. Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, direct delivery to the selected organ (e.g., intraportal delivery to the liver), oral, inhalation (including intranasal and intratracheal delivery), intraocular, intravenous, intramuscular, subcutaneous, intradermal, and other parental routes of administration. Routes of administration may be combined, if desired.

**[0130]** Dosages of the viral vector will depend primarily on factors such as the condition being treated, the age, weight and health of the patient, and may thus vary among patients. For example, a therapeutically effective human dosage of the viral vector is generally in the range of from about 1 ml to about 100 ml of solution containing concentrations of from about  $1 \times 10^9$  to  $1 \times 10^{16}$  genomes virus vector. A preferred human dosage may be about  $1 \times 10^{13}$  to  $1 \times 10^{16}$  AAV genomes. The dosage will be adjusted to balance the therapeutic benefit against any side effects and

such dosages may vary depending upon the therapeutic application for which the recombinant vector is employed. The levels of expression of the transgene can be monitored to determine the frequency of dosage resulting in viral vectors, preferably AAV vectors containing the minigene. Optionally, dosage regimens similar to those described for therapeutic purposes may be utilized for immunization using the compositions of the invention.

**[0131]** Examples of therapeutic products and immunogenic products for delivery by the AAV-containing vectors of the invention are provided below. These vectors may be used for a variety of therapeutic or vaccinal regimens, as described herein. Additionally, these vectors may be delivered in combination with one or more other vectors or active ingredients in a desired therapeutic and/or vaccinal regimen.

## B. Therapeutic Transgenes

**[0132]** Useful therapeutic products encoded by the transgene include hormones and growth and differentiation factors including, without limitation, insulin, glucagon, growth hormone (GH), parathyroid hormone (PTH), growth hormone releasing factor (GRF), follicle stimulating hormone (FSH), luteinizing hormone (LH), human chorionic gonadotropin (hCG), vascular endothelial growth factor (VEGF), angiopoietins, angiostatin, granulocyte colony stimulating factor (GCSF), erythropoietin (EPO), connective tissue growth factor (CTGF), basic fibroblast growth factor (bFGF), acidic fibroblast growth factor (aFGF), epidermal growth factor (EGF), transforming growth factor  $\alpha$  (TGF $\alpha$ ), platelet-derived growth factor (PDGF), insulin growth factors I and II (IGF-I and IGF-11), any one of the transforming growth factor  $\beta$  superfamily, including TGF  $\beta$ , activins, inhibins, or any of the bone morphogenic proteins (BMP) BMPs 1-15, any one of the heregluin/neuregulin/ARIA/neu differentiation factor (NDF) family of growth factors, nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophins NT-3 and NT-4/5, ciliary neurotrophic factor (CNTF), glial cell line derived neurotrophic factor (GDNF), neurturin, agrin, any one of the family of semaphorins/collapsins, netrin-1 and netrin-2, hepatocyte growth factor (HGF), ephrins, noggin, sonic hedgehog and tyrosine hydroxylase.

**[0133]** Other useful transgene products include proteins that regulate the immune system including, without limitation, cytokines and lymphokines such as thrombopoietin (TPO), interleukins (IL) IL-1 through IL-25 (including, IL-2, IL-4, IL-12, and IL-18), monocyte chemoattractant protein, leukemia inhibitory factor, granulocyte-macrophage colony stimulating factor, Fas ligand, tumor necrosis factors  $\alpha$  and  $\beta$ , interferons  $\alpha$ ,  $\beta$ , and  $\gamma$ , stem cell factor, flk-2/flt3 ligand. Gene products produced by the immune system are also useful in the invention. These include, without limitations, immunoglobulins IgG, IgM, IgA, IgD and IgE, chimeric immunoglobulins, humanized antibodies, single chain antibodies, T cell receptors, chimeric T cell receptors, single chain T cell receptors, class I and class II MHC molecules, as well as engineered immunoglobulins and MHC molecules. Useful gene products also include complement regulatory proteins such as complement regulatory proteins, membrane cofactor protein (MCP), decay accelerating factor (DAF), CR1, CF2 and CD59.

**[0134]** Still other useful gene products include any one of the receptors for the hormones, growth factors, cytokines, lymphokines, regulatory proteins and immune system proteins. The invention encompasses receptors for cholesterol regulation, including the low density lipoprotein (LDL) receptor, high density lipoprotein (HDL) receptor, the very low density lipoprotein (VLDL) receptor, and the scavenger receptor. The invention also encompasses gene products such as members of the steroid hormone receptor superfamily including glucocorticoid receptors and estrogen receptors, Vitamin D receptors and other nuclear receptors. In addition, useful gene products include transcription factors such as *jun*, *fos*, max, mad, serum response factor (SRF), AP-1, AP2, *myb*, MyoD and myogenin, ETS-box containing proteins, TFE3, E2F, ATF1, ATF2, ATF3, ATF4, ZF5, NFAT, CREB, HNF-4, C/EBP, SP1, CCAAT-box binding proteins, interferon regulation factor (IRF-1), Wilms tumor protein, ETS-binding protein, STAT, GATA-box binding proteins, e.g., GATA-3, and the forkhead family of winged helix proteins.

**[0135]** Other useful gene products include, carbamoyl synthetase 1, ornithine transcarbamylase, arginosuccinate synthetase, arginosuccinate lyase, arginase, fumarylacetic acid hydrolase, phenylalanine hydroxylase, alpha-1 anti-trypsin, glucose-6-phosphatase, porphobilinogen deaminase, factor VIII, factor IX, cystathione beta-synthase, branched chain ketoacid decarboxylase, albumin, isovaleryl-CoA dehydrogenase, propionyl CoA carboxylase, methyl malonyl CoA mutase, glutaryl CoA dehydrogenase, insulin, beta-glucosidase, pyruvate carboxylase, hepatic phosphorylase, phosphorylase kinase, glycine decarboxylase, H-protein, T-protein, a cystic fibrosis transmembrane regulator (CFTR) sequence, and a dystrophin cDNA sequence. Still other useful gene products include enzymes such as may be useful in enzyme replacement therapy, which is useful in a variety of conditions resulting from deficient activity of enzyme. For example, enzymes that contain mannose-6-phosphate may be utilized in therapies for lysosomal storage diseases (e.g., a suitable gene includes that encoding  $\beta$ -glucuronidase (GUSB)).

**[0136]** Other useful gene products include non-naturally occurring polypeptides, such as chimeric or hybrid polypeptides having a non-naturally occurring amino acid sequence containing insertions, deletions or amino acid substitutions. For example, single-chain engineered immunoglobulins could be useful in certain immunocompromised patients. Other types of non-naturally occurring gene sequences include antisense molecules and catalytic nucleic acids, such as ribozymes, which could be used to reduce overexpression of a target.

**[0137]** Reduction and/or modulation of expression of a gene is particularly desirable for treatment of hyperproliferative conditions characterized by hyperproliferating cells, as are cancers and psoriasis. Target polypeptides include those polypeptides which are produced exclusively or at higher levels in hyperproliferative cells as compared to normal cells. Target antigens include polypeptides encoded by oncogenes such as myb, myc, fyn, and the translocation gene bcr/abl, ras, src, P53, neu, trk and EGRF. In addition to oncogene products as target antigens, target polypeptides for anti-cancer treatments and protective regimens include variable regions of antibodies made by B cell lymphomas and variable regions of T cell receptors of T cell lymphomas which, in some embodiments, are also used as target antigens for autoimmune disease. Other tumor-associated polypeptides can be used as target polypeptides such as polypeptides which are found at higher levels in tumor cells including the polypeptide recognized by monoclonal antibody 17-1A and folate binding polypeptides.

**[0138]** Other suitable therapeutic polypeptides and proteins include those which may be useful for treating individuals suffering from autoimmune diseases and disorders by conferring a broad based protective immune response against targets that are associated with autoimmunity including cell receptors and cells which produce "self"-directed antibodies. T cell mediated autoimmune diseases include Rheumatoid arthritis (RA), multiple sclerosis (MS), Sjögren's syndrome, sarcoidosis, insulin dependent diabetes mellitus (IDDM), autoimmune thyroiditis, reactive arthritis, ankylosing spondylitis, scleroderma, polymyositis, dermatomyositis, psoriasis, vasculitis, Wegener's granulomatosis, Crohn's disease and ulcerative colitis. Each of these diseases is characterized by T cell receptors (TCRs) that bind to endogenous antigens and initiate the inflammatory cascade associated with autoimmune diseases.

### C. Immunogenic Transgenes

**[0139]** Alternatively, or in addition, the vectors of the invention may contain AAV sequences of the invention and a transgene encoding a peptide, polypeptide or protein which induces an immune response to a selected immunogen. For example, immunogens may be selected from a variety of viral families. Example of desirable viral families against which an immune response would be desirable include, the picornavirus family, which includes the genera rhinoviruses, which are responsible for about 50% of cases of the common cold; the genera enteroviruses, which include polioviruses, coxsackieviruses, echoviruses, and human enteroviruses such as hepatitis A virus; and the genera aphoviruses, which are responsible for foot and mouth diseases, primarily in non-human animals. Within the picornavirus family of viruses, target antigens include the VP1, VP2, VP3, VP4, and VPG. Another viral family includes the calcivirus family, which encompasses the Norwalk group of viruses, which are an important causative agent of epidemic gastroenteritis. Still another viral family desirable for use in targeting antigens for inducing immune responses in humans and non-human animals is the togavirus family, which includes the genera alphavirus, which include Sindbis viruses, Ross River virus, and Venezuelan, Eastern & Western Equine encephalitis, and rubivirus, including Rubella virus. The flaviviridae family includes dengue, yellow fever, Japanese encephalitis, St. Louis encephalitis and tick borne encephalitis viruses. Other target antigens may be generated from the Hepatitis C or the coronavirus family, which includes a number of non-human viruses such as infectious bronchitis virus (poultry), porcine transmissible gastroenteric virus (pig), porcine hemagglutinating encephalomyelitis virus (pig), feline infectious peritonitis virus (cats), feline enteric coronavirus (cat), canine coronavirus (dog), and human respiratory coronaviruses, which may cause the common cold and/or non-A, B or C hepatitis. Within the coronavirus family, target antigens include the E1 (also called M or matrix protein), E2 (also called S or Spike protein), E3 (also called HE or hemagglutinin-esterase) glycoprotein (not present in all coronaviruses), or N (nucleocapsid). Still other antigens may be targeted against the rhabdovirus family, which includes the genera vesiculovirus (e.g., Vesicular Stomatitis Virus), and the general lyssavirus (e.g., rabies). Within the rhabdovirus family, suitable antigens may be derived from the G protein or the N protein. The family filoviridae, which includes hemorrhagic fever viruses such as Marburg and Ebola virus may be a suitable source of antigens. The paramyxovirus family includes parainfluenza Virus Type 1, parainfluenza Virus Type 3, bovine parainfluenza Virus Type 3, rubulavirus (mumps virus, parainfluenza Virus Type 2, parainfluenza virus Type 4, Newcastle disease virus (chickens), rinderpest, morbillivirus, which includes measles and canine distemper, and pneumovirus, which includes respiratory syncytial virus. The influenza virus is classified within the family orthomyxovirus and is a suitable source of antigen (e.g., the HA protein, the N1 protein). The bunyavirus family includes the genera bunyavirus (California encephalitis, La Crosse), phlebovirus (Rift Valley Fever), hantavirus (pneumonia is a hemorrhagic fever virus),airovirus (Nairobi sheep disease) and various unassigned bunyaviruses. The arenavirus family provides a source of antigens against LCM and Lassa fever virus. The reovirus family includes the genera reovirus, rotavirus (which causes acute gastroenteritis in children), orbiviruses, and cultivirus (Colorado Tick fever, Lebombo (humans), equine encephalosis, blue tongue).

**[0140]** The retrovirus family includes the sub-family oncorivirinal which encompasses such human and veterinary diseases as feline leukemia virus, HTLV I and HTLV II, lentivirinal (which includes human immunodeficiency virus (HIV), simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV), equine infectious anemia virus, and spumavirinal). Between the HIV and SIV, many suitable antigens have been described and can readily be selected. Examples of suitable HIV and SIV antigens include, without limitation the gag, pol, Vif, Vpx, VPR, Env, Tat and Rev proteins, as

well as various fragments thereof. In addition, a variety of modifications to these antigens have been described. Suitable antigens for this purpose are known to those of skill in the art. For example, one may select a sequence encoding the gag, pol, Vif, and Vpr, Env, Tat and Rev, amongst other proteins. See, e.g., the modified gag protein which is described in US Patent 5,972,596. See, also, the HIV and SIV proteins described in D.H. Barouch et al, J. Virol., 75(5):2462-2467 (March 2001), and R.R. Amara, et al, Science, 292:69-74 (6 April 2001). These proteins or subunits thereof may be delivered alone, or in combination via separate vectors or from a single vector.

**[0141]** The papovavirus family includes the sub-family polyomaviruses (BKU and JCU viruses) and the sub-family papillomavirus (associated with cancers or malignant progression of papilloma). The adenovirus family includes viruses (EX, AD7, ARD, O.B.) which cause respiratory disease and/or enteritis. The parvovirus family feline parvovirus (feline enteritis), feline panleucopeniavirus, canine parvovirus, and porcine parvovirus. The herpesvirus family includes the sub-family alphaherpesvirinae, which encompasses the genera simplexvirus (HSV1, HSV2), varicellovirus (pseudorabies, varicella zoster) and the sub-family betaherpesvirinae, which includes the genera cytomegalovirus (HCMV, muromegalovirus) and the sub-family gammaherpesvirinae, which includes the genera lymphocryptovirus, EBV (Burkitt's lymphoma), infectious rhinotracheitis, Marek's disease virus, and rhadinovirus. The poxvirus family includes the sub-family chordopoxvirinae, which encompasses the genera orthopoxvirus (Variola (Smallpox) and Vaccinia (Cowpox)), parapoxvirus, avipoxvirus, capripoxvirus, leporipoxvirus, suipoxvirus, and the sub-family entomopoxvirinae. The hepadnavirus family includes the Hepatitis B virus. One unclassified virus which may be suitable source of antigens is the Hepatitis delta virus. Still other viral sources may include avian infectious bursal disease virus and porcine respiratory and reproductive syndrome virus. The alphavirus family includes equine arteritis virus and various Encephalitis viruses.

**[0142]** The present invention may also encompass immunogens which are useful to immunize a human or non-human animal against other pathogens including bacteria, fungi, parasitic microorganisms or multicellular parasites which infect human and non-human vertebrates, or from a cancer cell or tumor cell. Examples of bacterial pathogens include pathogenic gram-positive cocci include pneumococci; staphylococci; and streptococci. Pathogenic gram-negative cocci include meningococcus; gonococcus. Pathogenic enteric gram-negative bacilli include enterobacteriaceae; pseudomonas, acinetobacteria and eikenella; melioidosis; salmonella; shigella; haemophilus; moraxella; *H. ducreyi* (which causes chancroid); bruceila; *Francisella tularensis* (which causes tularemia); yersinia (pasteurella); streptobacillus moniliformis and spirillum; Gram-positive bacilli include listeria monocytogenes; erysipelothrix rhusiopathiae; *Corynebacterium diphtheria* (diphtheria); cholera; *B. anthracis* (anthrax); donovanosis (granuloma inguinale); and bartonellosis. Diseases caused by pathogenic anaerobic bacteria include tetanus; botulism; other clostridia; tuberculosis; leprosy; and other mycobacteria. Pathogenic spirochetal diseases include syphilis; treponematoses: yaws, pinta and endemic syphilis; and leptospirosis. Other infections caused by higher pathogen bacteria and pathogenic fungi include actinomycosis; nocardiosis; cryptococcosis, blastomycosis, histoplasmosis and coccidioidomycosis; candidiasis, aspergillosis, and mucormycosis; sporotrichosis; paracoccidioidomycosis, petriellidiosis, torulopsosis, mycetoma and chromomycosis; and dermatophytosis. Rickettsial infections include Typhus fever, Rocky Mountain spotted fever, Q fever, and Rickettsialpox. Examples of mycoplasma and chlamydial infections include: mycoplasma pneumoniae; lymphogranuloma venereum; psittacosis; and perinatal chlamydial infections. Pathogenic eukaryotes encompass pathogenic protozoans and helminths and infections produced thereby include: amebiasis; malaria; leishmaniasis; trypanosomiasis; toxoplasmosis; *Pneumocystis carinii*; *Trichans*; *Toxoplasma gondii*; babesiosis; giardiasis; trichinosis; filariasis; schistosomiasis; nematodes; trematodes or flukes; and cestode (tapeworm) infections.

**[0143]** Many of these organisms and/or toxins produced thereby have been identified by the Centers for Disease Control [(CDC), Department of Health and Human Services, USA], as agents which have potential for use in biological attacks. For example, some of these biological agents, include, *Bacillus anthracis* (anthrax), *Clostridium botulinum* and its toxin (botulism), *Yersinia pestis* (plague), variola major (smallpox), *Francisella tularensis* (tularemia), and viral hemorrhagic fever, all of which are currently classified as Category A agents; *Coxiella burnetii* (Q fever); Brucella species (brucellosis), *Burkholderia mallei* (glanders), *Ricinus communis* and its toxin (ricin toxin), *Clostridium perfringens* and its toxin (epsilon toxin), *Staphylococcus* species and their toxins (enterotoxin B), all of which are currently classified as Category B agents; and Nipah virus and hantaviruses, which are currently classified as Category C agents. In addition, other organisms, which are so classified or differently classified, may be identified and/or used for such a purpose in the future. It will be readily understood that the viral vectors and other constructs described herein are useful to deliver antigens from these organisms, viruses, their toxins or other byproducts, which will prevent and/or treat infection or other adverse reactions with these biological agents.

**[0144]** Administration of the vectors of the invention to deliver immunogens against the variable region of the T cells elicit an immune response including CTLs to eliminate those T cells. In rheumatoid arthritis (RA), several specific variable regions of T cell receptors (TCRs) which are involved in the disease have been characterized. These TCRs include V-3, V-14, V-17 and V $\alpha$ -17. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in RA. In multiple sclerosis (MS), several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-

7 and V $\alpha$ -10. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in MS. In scleroderma, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-6, V-8, V-14 and V $\alpha$ -16, V $\alpha$ -3C, V $\alpha$ -7, V $\alpha$ -14, V $\alpha$ -15, V $\alpha$ -16, V $\alpha$ -28 and V $\alpha$ -12. Thus, delivery of a nucleic acid molecule that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in scleroderma.

**[0145]** Optionally, vectors containing AAV sequences of the invention may be delivered using a prime-boost regimen. A variety of such regimens have been described in the art and may be readily selected. See, e.g., WO 00/11140, published March 2, 2000, incorporated by reference.

**[0146]** Such prime-boost regimens typically involve the administration of a DNA (e.g., plasmid) based vector to prime the immune system to second, booster, administration with a traditional antigen, such as a protein or a recombinant virus carrying the sequences encoding such an antigen. In one embodiment, the invention provides a method of priming and boosting an immune response to a selected antigen by delivering a plasmid DNA vector carrying said antigen, followed by boosting, e.g., with a vector containing AAV sequences of the invention.

**[0147]** In one embodiment, the prime-boost regimen involves the expression of multiproteins from the prime and/or the boost vehicle. See, e.g., R.R. Amara, Science, 292:69-74 (6 April 2001) which describes a multiprotein regimen for expression of protein subunits useful for generating an immune response against HIV and SIV. For example, a DNA prime may deliver the Gag, Pol, Vif, VPX and Vpr and Env, Tat, and Rev from a single transcript. Alternatively, the SIV Gag, Pol and HIV-1 Env is delivered.

**[0148]** However, the prime-boost regimens are not limited to immunization for HIV or to delivery of these antigens. For example, priming may involve delivering with a first chimp vector of the invention followed by boosting with a second chimp vector, or with a composition containing the antigen itself in protein form. In one or example, the prime-boost regimen can provide a protective immune response to the virus, bacteria or other organism from which the antigen is derived. In another desired embodiment, the prime-boost regimen provides a therapeutic effect that can be measured using convention assays for detection of the presence of the condition for which therapy is being administered.

**[0149]** The priming vaccine may be administered at various sites in the body in a dose dependent manner, which depends on the antigen to which the desired immune response is being targeted. The invention is not limited to the amount or situs of injection(s) or to the pharmaceutical carrier. Rather, the priming step encompasses treatment regimens which include a single dose or dosage which is administered hourly, daily, weekly or monthly, or yearly. As an example, the mammals may receive one or two priming injection containing between about 10  $\mu$ g to about 50  $\mu$ g of plasmid in carrier. A desirable priming amount or dosage of the priming DNA vaccine composition ranges between about 1  $\mu$ g to about 10,000  $\mu$ g of the DNA vaccine. Dosages may vary from about 1  $\mu$ g to 1000  $\mu$ g DNA per kg of subject body weight. The amount or site of injection is desirably selected based upon the identity and condition of the mammal being vaccinated.

**[0150]** The dosage unit of the DNA vaccine suitable for delivery of the antigen to the mammal is described herein. The DNA vaccine is prepared for administration by being suspended or dissolved in a pharmaceutically or physiologically acceptable carrier such as isotonic saline, isotonic salts solution or other formulations which will be apparent to those skilled in such administration. The appropriate carrier will be evident to those skilled in the art and will depend in large part upon the route of administration. The compositions of the invention may be administered to a mammal according to the routes described above, in a sustained release formulation using a biodegradable biocompatible polymer, or by on-site delivery using micelles, gels and liposomes.

**[0151]** Optionally, the priming step of this invention also includes administering with the priming DNA vaccine composition, a suitable amount of an adjuvant, such as are defined herein.

**[0152]** Preferably, a boosting composition is administered about 2 to about 27 weeks after administering the priming DNA vaccine to the mammalian subject. The administration of the boosting composition is accomplished using an effective amount of a boosting vaccine composition containing or capable of delivering the same antigen as administered by the priming DNA vaccine. The boosting composition may be composed of a recombinant viral vector derived from the same viral source or from another source. Alternatively, the "boosting composition" can be a composition containing the same antigen as encoded in the priming DNA vaccine, but in the form of a protein or peptide, which composition induces an immune response in the host. In another embodiment, the boosting vaccine composition includes a composition containing a DNA sequence encoding the antigen under the control of a regulatory sequence directing its expression in a mammalian cell, e.g., vectors such as well-known bacterial or viral vectors. The primary requirements of the boosting vaccine composition are that the antigen of the vaccine composition is the same antigen, or a cross-reactive antigen, as that encoded by the DNA vaccine.

**[0153]** Suitably, the vectors of the invention are also well suited for use in regimens which use non-AAV vectors as well as proteins, peptides, and/or other biologically useful therapeutic or immunogenic compounds. These regimens are particularly well suited to gene delivery for therapeutic poses and for immunization, including inducing protective immunity. Such uses will be readily apparent to one of skill in the art.

**[0154]** Further, a vector of the invention provides an efficient gene transfer vehicle which can deliver a selected

transgene to a selected host cell *in vivo or ex vivo* even where the organism has neutralizing antibodies to one or more AAV serotypes. In one embodiment, the vector (e.g., an rAAV) and the cells are mixed *ex vivo*; the infected cells are cultured using conventional methodologies; and the transduced cells are re-infused into the patient. Further, the vectors of the invention may also be used for production of a desired gene product *in vitro*. For *in vitro* production, a desired product (e.g., a protein) may be obtained from a desired culture following transfection of host cells with a rAAV containing the molecule encoding the desired product and culturing the cell culture under conditions which permit expression. The expressed product may then be purified and isolated, as desired. Suitable techniques for transfection, cell culturing, purification, and isolation are known to those of skill in the art.

**[0155]** The following examples illustrate several aspects and embodiments of the invention.

## EXAMPLES

Example 1: PCR amplification, cloning and characterization of novel AAV sequences.

**[0156]** Tissues from nonhuman primates were screened for AAV sequences using a PCR method based on oligonucleotides to highly conserved regions of known AAVs. A stretch of AAV sequence spanning 2886 to 3143 bp of AAV1 [SEQ ID NO:6] was selected as a PCR amplicon in which a hypervariable region of the capsid protein (Cap) that is unique to each known AAV serotype, which is termed herein a "signature region," is flanked by conserved sequences. In later analysis, this signature region was shown to be located between conserved residues spanning hypervariable region 3.

**[0157]** An initial survey of peripheral blood of a number of nonhuman primate species revealed detectable AAV in a subset of animals from species such as rhesus macaques, cynomolgous macaques, chimpanzees and baboons. However, there were no AAV sequences detected in some other species tested, including Japanese macaques, pig-tailed macaques and squirrel monkeys. A more extensive analysis of vector distribution was conducted in tissues of rhesus monkeys of the University of Pennsylvania and Tulane colonies recovered at necropsy. This revealed AAV sequence throughout a wide array of tissues.

### A. Amplification of an AAV signature region

**[0158]** DNA sequences of AAV1-6 and AAVs isolated from Goose and Duck were aligned to each other using "Clustal W" at default settings. The alignment for AAV1-6, and including the information for the novel AAV7, is provided in Fig. 1. Sequence similarities among AAVs were compared.

**[0159]** In the line of study, a 257 bp region spanning 2886 bp to 3143 bp of AAV 1 [SEQ ID NO: 6], and the corresponding region in the genomes of AAV 2-6 genomes [See, Fig. 1], was identified by the inventors. This region is located with the AAV capsid gene and has highly conserved sequences among at both 5' and 3' ends and is relatively variable sequence in the middle. In addition, this region contains a DraIII restriction enzyme site (CACCACGTC, SEQ ID NO:15). The inventors have found that this region serves as specific signature for each known type of AAV DNA. In other words, following PCR reactions, digestion with endonucleases that are specific to each known serotypes and gel electrophoresis analysis, this regions can be used to definitively identify amplified DNA as being from serotype 1, 2, 3, 4, 5, 6, or another serotype.

**[0160]** The primers were designed, validated and PCR conditions optimized with AAV1, 2 and 5 DNA controls. The primers were based upon the sequences of AAV2: 5' primer, 1S: bp 2867-2891 of AAV2 (SEQ ID NO:7) and 3' primer, 18as, bp 3095-3121 of AAV2 (SEQ ID NO:7).

**[0161]** Cellular DNAs from different tissues including blood, brain, liver, lung, testis, etc. of different rhesus monkeys were studied utilizing the strategy described above. The results revealed that DNAs from different tissues of these monkeys gave rise to strong PCR amplifications. Further restriction analyses of PCR products indicated that they were amplified from AAV sequences different from any published AAV sequences.

**[0162]** PCR products (about 255 bp in size) from DNAs of a variety of monkey tissues have been cloned and sequenced. Bioinformatics study of these novel AAV sequences indicated that they are novel AAV sequences of capsid gene and distinct from each other. Fig. 1 includes in the alignment the novel AAV signature regions for AAV10-12 [SEQ ID NO:117, 118 and 119, respectively]. Multiple sequence alignment analysis was performed using the Clustal W (1.81) program. The percentage of sequence identity between the signature regions of AAV 1-7 and AAV 10-12 genomes is provided below.

Table 1.

Sequences for Analysis		
Sequence #	AAV Serotype	Size (bp)
1	AAV1	258
2	AAV2	255
3	AAV3	255
4	AAV4	246
5	AAV5	258
6	AAV6	258
7	AAV7	258
10	AAV10	255
11	AAV11	258
12	AAV12	255

Table 3.

Pairwise Alignment (Percentage of Identity)									
	AAV2	AAV3	AAV4	AAV5	AAV6	AAV7	AAV10	AAV11	AAV12
AAV1	90	90	81	76	97	91	93	94	93
AAV2		93	79	78	90	90	93	93	92
AAV3			80	76	90	92	92	92	92
AAV4				76	81	84	82	81	79
AAV5					75	78	79	79	76
AAV6						91	92	94	94
AAV7							94	92	92
AAV10								95	93
AAV11									94

**[0163]** Over 300 clones containing novel AAV serotype sequences that span the selected 257 bp region were isolated and sequenced. Bioinformatics analysis of these 300+ clones suggests that this 257 bp region is critical in serving as a good land marker or signature sequence for quick isolation and identification of novel AAV serotype.

*B. Use of the signature region for PCR amplification.*

**[0164]** The 257 bp signature region was used as a PCR anchor to extend PCR amplifications to 5' of the genome to cover the junction region of rep and cap genes (1398 bp - 3143 bp, SEQ ID NO:6) and 3' of the genome to obtain the entire cap gene sequence (2866 bp - 4600 bp, SEQ ID NO:6). PCR amplifications were carried out using the standard conditions, including denaturing at 95°C for 0.5-1 min, annealing at 60-65°C for 0.5-1 min and extension at 72° C for 1 min per kb with a total number of amplification cycles ranging from 28 to 42.

**[0165]** Using the aligned sequences as described in "A", two other relative conserved regions were identified in the sequence located in 3' end of rep genes and 5' to the 257 bp region and in the sequence down stream of the 257 bp fragment but before the AAV' 3 ITR. Two sets of new primers were designed and PCR conditions optimized for recovery of entire capsid and a part of rep sequences of novel AAV serotypes. More specifically, for the 5' amplification, the 5' primer, AV1Ns, was GCTGCGTCAACTGGACCAATGAGAAC [nt 1398-1423 of AAV1, SEQ ID NO:6] and the 3' primer

was 18as, identified above. For the 3' amplification, the 5' primer was 1s, identified above, and the 3' primer was AV2Las, TCGTTTCAGTTGAACTTTGGTCTCTGCG [nt 4435-4462 of AAV2, SEQ ID NO:7].

**[0166]** In these PCR amplifications, the 257 bp region was used as a PCR anchor and land marker to generate overlapping fragments to construct a complete capsid gene by fusion at the DraIII site in the signature region following amplification of the 5' and 3' extension fragments obtained as described herein. More particularly, to generate the intact AAV7 cap gene, the three amplification products (a) the sequences of the signature region; (b) the sequences of the 5' extension; and (c) the sequences of the 3' extension were cloned into a pCR4-Topo [Invitrogen] plasmid backbone according to manufacturer's instructions. Thereafter, the plasmids were digested with DraIII and recombined to form an intact cap gene.

**[0167]** In this line of work, about 80 % of capsid sequences of AAV7 and AAV 8 were isolated and analyzed. Another novel serotype, AAV9, was also discovered from Monkey #2.

**[0168]** Using the PCR conditions described above, the remaining portion of the rep gene sequence for AAV7 is isolated and cloned using the primers that amplify 108 bp to 1461 bp of AAV genome (calculated based on the numbering of AAV2, SEQ ID NO:7). This clone is sequenced for construction of a complete AAV7 genome without ITRs.

### C. Direct Amplification of 3.1 kb Cap fragment

**[0169]** To directly amplify a 3.1 kb full-length Cap fragment from NHP tissue and blood DNAs, two other highly conserved regions were identified in AAV genomes for use in PCR amplification of large fragments. A primer within a conserved region located in the middle of the rep gene was selected (AV1ns: 5' GCTGCGTCAACTGGACCAATGA-GAAC 3', nt 1398-1423 of SEQ ID NO:6) in combination with the 3' primer located in another conserved region downstream of the Cap gene (AV2cas: 5' CGCAGAGACCAAAGTTCAACTGAAACGA 3', SEQ ID NO:7) for amplification of full-length cap fragments. The PCR products were Topo-cloned according to manufacturer's directions (Invitrogen) and sequence analysis was performed by Qiagen genomics (Qiagen genomics, Seattle, WA) with an accuracy of  $\geq 99.9\%$ . A total of 50 capsid clones were isolated and characterized. Among them, 37 clones were derived from Rhesus macaque tissues (rh.1 - rh.37), 6 clones from cynomolgous macaques (cy.1 - cy.6), 2 clones from Baboons (bb.1 and bb.2) and 5 clones from Chimps (ch.1 - ch.5).

**[0170]** To rule out the possibility that sequence diversity within the novel AAV family was not an artifact of the PCR, such as PCR-mediated gene splicing by overlap extension between different partial DNA templates with homologous sequences, or the result of recombination process in bacteria, a series of experiments were performed under identical conditions for VP1 amplification using total cellular DNAs. First, intact AAV7 and AAV8 plasmids were mixed at an equal molar ratio followed by serial dilutions. The serially diluted mixtures were used as templates for PCR amplification of 3.1 kb VP1 fragments using universal primers and identical PCR conditions to that were used for DNA amplifications to see whether any hybrid PCR products were generated. The mixture was transformed into bacteria and isolated transformants to look for hybrid clones possibly derived from recombination process in bacterial cells. In a different experiment, we restricted AAV7 and AAV8 plasmids with Msp I, Ava I and HaeI, all of which cut both genomes multiple times at different positions, mixed the digestions in different combinations and used them for PCR amplification of VP1 fragments under the same conditions to test whether any PCR products could be generated through overlap sequence extension of partial AAV sequences. In another experiment, a mixture of gel purified 5' 1.5 kb AAV7 VP1 fragment and 3' 1.7 kb AAV8 VP1 fragment with overlap in the signature region was serially diluted and used for PCR amplification in the presence and absence of 200 ng cellular DNA extracted from a monkey cell line that was free of AAV sequences by TaqMan analysis. None of these experiments demonstrated efficient PCR-mediated overlap sequence production under the conditions of the genomic DNA Cap amplification (data not shown). As a further confirmation, 3 pairs of primers were designed, which were located at different HVRs, and were sequence specific to the variants of clone 42s from Rhesus macaque F953, in different combinations to amplify shorter fragments from mesenteric lymph node (MLN) DNA from F953 from which clone 42s were isolated. All sequence variations identified in full-length Cap clones were found in these short fragments (data not shown).

### Example 2: Adeno-Associated Viruses Undergo Substantial Evolution in Primates During Natural Infections

**[0171]** Sequence analysis of selected AAV isolates revealed divergence throughout the genome that is most concentrated in hypervariable regions of the capsid proteins. Epidemiologic data indicate that all known serotypes are endemic to primates, although isolation of clinical isolates has been restricted to AAV2 and AAV3 from anal and throat swabs of human infants and AAV5 from a human condylomatous wart. No known clinical sequelae have been associated with AAV infection.

**[0172]** In an attempt to better understand the biology of AAV, nonhuman primates were used as models to characterize the sequelae of natural infections. Tissues from nonhuman primates were screened for AAV sequences using the PCR method of the invention based on oligonucleotides to highly conserved regions of known AAVs (see Example 1).

A stretch of AAV sequence spanning 2886 to 3143 bp of AAV1 [SEQ ID NO:6] was selected as a PCR amplicon in which conserved sequences are flanked by a hypervariable region that is unique to each known AAV serotype, termed herein a "signature region."

**[0173]** An initial survey of peripheral blood of a number of nonhuman primate species including rhesus monkeys, cynomolgous monkeys, chimpanzees, and baboons revealed detectable AAV in a subset of animals from all species. A more extensive analysis of vector distribution was conducted in tissues of rhesus monkeys of the University of Pennsylvania and Tulane colonies recovered at necropsy. This revealed AAV sequence throughout a wide array of tissues.

**[0174]** The amplified signature sequences were subcloned into plasmids and individual transformants were subjected to sequence analysis. This revealed substantial variation in nucleotide sequence of clones derived from different animals. Variation in the signature sequence was also noted in clones obtained within individual animals. Tissues harvested from two animals in which unique signature sequences were identified (i.e., colon from 98E044 and heart from 98E056) were further characterized by expanding the sequence amplified by PCR using oligonucleotides to highly conserved sequences. In this way, complete proviral structures were reconstructed for viral genomes from both tissues as described herein. These proviruses differ from the other known AAVs with the greatest sequence divergence noted in regions of the Cap gene.

**[0175]** Additional experiments were performed to confirm that AAV sequences resident to the nonhuman primate tissue represented proviral genomes of infectious virus that is capable of being rescued and form virions. Genomic DNA from liver tissue of animal 98E056, from which AAV8 signature sequence was detected, was digested with an endonuclease that does not have a site within the AAV sequence and transfected into 293 cells with a plasmid containing an E1 deleted genome of human adenovirus serotype 5 as a source of helper functions. The resulting lysate was passaged on 293 cells once and the lysate was recovered and analyzed for the presence of AAV Cap proteins using a broadly reacting polyclonal antibody to Cap proteins and for the presence and abundance of DNA sequences from the PCR amplified AAV provirus from which AAV8 was derived. Transfection of endonuclease restricted heart DNA and the adenovirus helper plasmid yielded high quantities of AAV8 virus as demonstrated by the detection of Cap proteins by Western blot analysis and the presence of  $10^4$  AAV8 vector genomes per 293 cell. Lysates were generated from a large-scale preparation and the AAV was purified by cesium sedimentation. The purified preparation demonstrated 26 nm icosahedral structures that look identical to those of AAV serotype 2. Transfection with the adenovirus helper alone did not yield AAV proteins or genomes, ruling out contamination as a source of the rescued AAV.

**[0176]** To further characterize the inter and intra animal variation of AAV signature sequence, selected tissues were subjected to extended PCR to amplify entire Cap open reading frames.

**[0177]** The resulting fragments were cloned into bacterial plasmids and individual transformants were isolated and fully sequenced. This analysis involved mesenteric lymph nodes from three rhesus monkeys (Tulane/V223 - 6 clones; Tulane/T612 - 7 clones; Tulane/F953 - 14 clones), liver from two rhesus monkeys (Tulane/V251 - 3 clones; Penn/00E033 - 3 clones), spleen from one rhesus monkey (Penn/97E043 - 3 clones), heart from one rhesus monkey (IHGT/98E046 - 1 clone) and peripheral blood from one chimpanzee (New Iberia/X133 - 5 clones), six cynomolgous macaques (Charles River/A1378, A3099, A3388, A3442, A2821, A3242 - 6 clones total) and one Baboon (SFRB/8644 - 2 clones). Of the 50 clones that were sequenced from 15 different animals, 30 were considered non-redundant based on the finding of at least 7 amino acid differences from one another. The non-redundant VP1 clones are numbered sequentially as they were isolated, with a prefix indicating the species of non-human primate from which they were derived. The structural relationships between these 30 non-redundant clones and the previously described 8 AAV serotypes were determined using the SplitsTree program [Huson, D. H. SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* 14, 68-73 (1998)] with implementation of the method of split decomposition. The analysis depicts homoplasy between a set of sequences in a tree-like network rather than a bifurcating tree. The advantage is to enable detection of groupings that are the result of convergence and to exhibit phylogenetic relationships even when they are distorted by parallel events. Extensive phylogenetic research will be required in order to elucidate the AAV evolution, whereas the intention here only is to group the different clones as to their sequence similarity.

**[0178]** To confirm that the novel VP1 sequences were derived from infectious viral genomes, cellular DNA from tissues with high abundance of viral DNA was restricted with an endonuclease that should not cleave within AAV and transfected into 293 cells, followed by infection with adenovirus. This resulted in rescue and amplification of AAV genomes from DNA of tissues from two different animals (data not shown).

**[0179]** VP1 sequences of the novel AAVs were further characterized with respect to the nature and location of amino acid sequence variation. All 30 VP1 clones that were shown to differ from one another by greater than 1% amino acid sequence were aligned and scored for variation at each residue. An algorithm developed to determine areas of sequence divergence yielded 12 hypervariable regions (HVR) of which 5 overlap or are part of the 4 previously described variable regions [Kotin, cited above; Rutledge, cited above]. The three-fold-proximal peaks contain most of the variability (HVR5-10). Interestingly the loops located at the 2 and 5 fold axis show intense variation as well. The HVRs 1 and 2 occur in the N-terminal portion of the capsid protein that is not resolved in the X-ray structure suggesting that the N-terminus of the VP1 protein is exposed on the surface of the virion.

[0180] Real-time PCR was used to quantify AAV sequences from tissues of 21 rhesus monkeys using primers and probes to highly conserved regions of Rep (one set) and Cap (two sets) of known AAVs. Each data point represents analysis from tissue DNA from an individual animal. This confirmed the wide distribution of AAV sequences, although the quantitative distribution differed between individual animals. The source of animals and previous history or treatments did not appear to influence distribution of AAV sequences in rhesus macaques. The three different sets of primers and probes used to quantify AAV yielded consistent results. The highest levels of AAV were found consistently in mesenteric lymph nodes at an average of 0.01 copies per diploid genome for 13 animals that were positive. Liver and spleen also contained high abundance of virus DNA. There were examples of very high AAV, such as in heart of rhesus macaque 98E056, spleen of rhesus macaque 97E043 and liver of rhesus macaque RQ4407, which demonstrated 1.5, 3 and 20 copies of AAV sequence per diploid genome respectively. Relatively low levels of virus DNA were noted in peripheral blood mononuclear cells, suggesting the data in tissue are not due to resident blood components (data not shown). It should be noted that this method would not necessarily capture all AAVs resident to the nonhuman primates since detection requires high homology to both the oligonucleotides and the real time PCR probe. Tissues from animals with high abundance AAV DNA was further analyzed for the molecular state of the DNA, by DNA hybridization techniques, and its cellular distribution, by *in situ* hybridization.

[0181] The kind of sequence variation revealed in AAV proviral fragments isolated from different animals and within tissues of the same animals is reminiscent of the evolution that occurs for many RNA viruses during pandemics or even within the infection of an individual. In some situations the notion of a wild-type virus has been replaced by the existence of swarms of quasispecies that evolve as a result of rapid replication and mutations in the presence of selective pressure. One example is infection by HIV, which evolves in response to immunologic and pharmacologic pressure. Several mechanisms contribute to the high rate of mutations in RNA viruses, including low fidelity and lack of proof reading capacity of reverse transcriptase and non-homologous and homologous recombination.

[0182] Evidence for the formation of quasispecies of AAV was illustrated in this study by the systematic sequencing of multiple cloned proviral fragments. In fact, identical sequences could not be found within any extended clones isolated between or within animals. An important mechanism for this evolution of sequence appears to be a high rate of homologous recombination between a more limited number of parenteral viruses. The net result is extensive swapping of hypervariable regions of the Cap protein leading to an array of chimeras that could have different tropisms and serologic specificities (i.e., the ability to escape immunologic responses especially as it relates to neutralizing antibodies). Mechanisms by which homologous recombination could occur are unclear. One possibility is that + and - strands of different single stranded AAV genomes anneal during replication as has been described during high multiplicity of infections with AAV recombinants. It is unclear if other mechanisms contribute to sequence evolution in AAV infections. The overall rate of mutation that occurs during AAV replication appears to be relatively low and the data do not suggest high frequencies of replication errors. However, substantial rearrangements of the AAV genome have been described during lytic infection leading to the formation of defective interfering particles. Irrespective of the mechanisms that lead to sequence divergence, with few exceptions, vp1 structures of the quasispecies remained intact without frameshifts or nonsense mutations suggesting that competitive selection of viruses with the most favorable profile of fitness contribute to the population dynamics.

[0183] These studies have implications in several areas of biology and medicine. The concept of rapid virus evolution, formerly thought to be a property restricted to RNA viruses, should be considered in DNA viruses, which classically have been characterized by serologic assays. It will be important in terms of parvoviruses to develop a new method for describing virus isolates that captures the complexity of its structure and biology, such as with HIV, which are categorized as general families of similar structure and function called Clades. An alternative strategy is to continue to categorize isolates with respect to serologic specificity and develop criteria for describing variants within serologic groups.

Example 3: Vectorology of recombinant AAV genomes equipped with AAV2 ITRs using chimeric plasmids containing AAV2 rep and novel AAV cap genes for serological and gene transfer studies in different animal models.

[0184] Chimeric packaging constructs are generated by fusing AAV2 rep with cap sequences of novel AAV serotypes. These chimeric packaging constructs are used, initially, for pseudotyping recombinant AAV genomes carrying AAV2 ITRs by triple transfection in 293 cell using Ad5 helper plasmid. These pseudotyped vectors are used to evaluate performance in transduction-based serological studies and evaluate gene transfer efficiency of novel AAV serotypes in different animal models including NHP and rodents, before intact and infectious viruses of these novel serotypes are isolated.

A. pAAV2GFP

[0185] The AAV2 plasmid which contains the AAV2 ITRs and green fluorescent protein expressed under the control

of a constitutive promoter. This plasmid contains the following elements: the AAV2 ITRs, a CMV promoter, and the GFP coding sequences.

#### B. Cloning of trans plasmid

**[0186]** To construct the chimeric trans-plasmid for production of recombinant pseudotyped AAV7 vectors, p5E18 plasmid (Xiao *et al.*, 1999, *J. Virol* **73**:3994-4003) was partially digested with Xho I to linearize the plasmid at the Xho I site at the position of 3169 bp only. The Xho I cut ends were then filled in and ligated back. This modified p5E18 plasmid was restricted with Xba I and Xho I in a complete digestion to remove the AAV2 cap gene sequence and replaced with a 2267 bp Spe I/Xho I fragment containing the AAV7 cap gene which was isolated from pCRAAV7 6-5+15-4 plasmid.

**[0187]** The resulting plasmid contains the AAV2 rep sequences for Rep78/68 under the control of the AAV2 P5 promoter, and the AAV2 rep sequences for Rep52/40 under the control of the AAV2 P19 promoter. The AAV7 capsid sequences are under the control of the AAV2 P40 promoter, which is located within the Rep sequences. This plasmid further contains a spacer 5' of the rep ORF.

#### C. Production of Pseudotyped rAAV

**[0188]** The rAAV particles (AAV2 vector in AAV7 capsid) are generated using an adenovirus-free method. Briefly, the cis plasmid (pAAV2.1 lacZ plasmid containing AAV2 ITRs), and the trans plasmid pCRAAV7 6-5+15-4 (containing the AAV2 rep and AAV7 cap) and a helper plasmid, respectively, were simultaneously co-transfected into 293 cells in a ratio of 1:1:2 by calcium phosphate precipitation.

**[0189]** For the construction of the pAd helper plasmids, pBG10 plasmid was purchased from Microbix (Canada). A RsrII fragment containing L2 and L3 was deleted from pBHG10, resulting in the first helper plasmid, pAdΔF13. Plasmid AdΔ F1 was constructed by cloning Asp700/SalI fragment with a PmeI/SgfI deletion, isolating from pBHG10, into Bluescript. MLP, L2, L2 and L3 were deleted in the pAdΔF1. Further deletions of a 2.3 kb NruI fragment and, subsequently, a 0.5 kb RsrII/NruI fragment generated helper plasmids pAdΔF5 and pAdΔF6, respectively. The helper plasmid, termed pΔF6, provides the essential helper functions of E2a and E4 ORF6 not provided by the E1-expressing helper cell, but is deleted of adenoviral capsid proteins and functional E1 regions).

**[0190]** Typically, 50 µg of DNA (cis:trans:helper) was transfected onto a 150 mm tissue culture dish. The 293 cells were harvested 72 hours post-transfection, sonicated and treated with 0.5% sodium deoxycholate (37°C for 10 min.) Cell lysates were then subjected to two rounds of a CsCl gradient. Peak fractions containing rAAV vector are collected, pooled and dialyzed against PBS.

Example 4: Creation of infectious clones carrying intact novel AAV serotypes for study of basic virology in human and NHP derived cell lines and evaluation of pathogenesis of novel AAV serotypes in NHP and other animal models.

**[0191]** To achieve this goal, the genome walker system is employed to obtain 5' and 3' terminal sequences (ITRs) and complete construction of clones containing intact novel AAV serotype genomes.

**[0192]** Briefly, utilizing a commercially available Universal Genome Walker Kit [Clontech], genomic DNAs from monkey tissues or cell lines that are identified as positive for the presence of AAV7 sequence are digested with Dra I, EcoRV, Pvu II and Stu I endonucleases and ligated to Genome Walker Adaptor to generate 4 individual Genome Walker Libraries (GWLs). Using DNAs from GWLs as templates, AAV7 and adjacent genomic sequences will be PCR-amplified by the adaptor primer 1 (API, provided in the kit) and an AAV7 specific primer 1, followed by a nested PCR using the adaptor primer 2 (AP2) and another AAV7 specific primer 2, both of which are internal to the first set of primers. The major PCR products from the nested PCR are cloned and characterized by sequencing analysis.

**[0193]** In this experiment, the primers covering the 257 bp or other signature fragment of a generic AAV genome are used for PCR amplification of cellular DNAs extracted from Human and NHP derived cell lines to identify and characterize latent AAV sequences. The identified latent AAV genomes are rescued from the positive cell lines using adenovirus helpers of different species and strains.

**[0194]** To isolate infectious AAV clones from NHP derived cell lines, a desired cell line is obtained from ATCC and screened by PCR to identify the 257 bp amplicon, i.e., signature region of the invention. The 257 bp PCR product is cloned and serotyped by sequencing analysis. For these cell lines containing the AAV7 sequence, the cells are infected with SV-15, a simian adenovirus purchased from ATCC, human Ad5 or transfected with plasmid construct housing the human Ad genes that are responsible for AAV helper functions. At 48 hour post infection or transfection, the cells are harvested and Hirt DNA is prepared for cloning of AAV7 genome following Xiao *et al.*, 1999, *J. Virol*, **73**:3994-4003.

## Example 5 - Production of AAV Vectors

[0195] A pseudotyping strategy similar to that of Example 3 for AAV1/7 was employed to produce AAV2 vectors packaged with AAV1, AAV5 and AAV8 capsid proteins. Briefly, recombinant AAV genomes equipped with AAV2 ITRs were packaged by triple transfection of 293 cells with cis-plasmid, adenovirus helper plasmid and a chimeric packaging construct where the AAV2 rep gene is fused with cap genes of novel AAV serotypes. To create the chimeric packaging constructs, the Xho I site of p5E18 plasmid at 3169 bp was ablated and the modified plasmid was restricted with Xba I and Xho I in a complete digestion to remove the AAV2 cap gene and replace it with a 2267 bp Spe I/Xho I fragment containing the AAV8 cap gene [Xiao, W., et al., (1999) *J Virol* **73**, 3994-4003]. A similar cloning strategy was used for creation of chimeric packaging plasmids of AAV2/1 and AAV2/5. All recombinant vectors were purified by the standard CsCl<sub>2</sub> sedimentation method except for AAV2/2, which was purified by single step heparin chromatography.

[0196] Genome copy (GC) titers of AAV vectors were determined by TaqMan analysis using probes and primers targeting SV40 poly A region as described previously [Gao, G., et al., (2000) *Hum Gene Ther* **11**, 2079-91].

[0197] Vectors were constructed for each serotype for a number of *in vitro* and *in vivo* studies. Eight different transgene cassettes were incorporated into the vectors and recombinant virions were produced for each serotype. The recovery of virus, based on genome copies, is summarized in Table 4 below. The yields of vector were high for each serotype with no consistent differences between serotypes. Data presented in the table are average genome copy yields with standard deviation x 10<sup>13</sup> of multiple production lots of 50 plate (150 mm) transfections.

Table 4.

Production of Recombinant Vectors					
	AAV2/1	AAV2/2	AAV2/5	AAV2/7	AAV2/8
<b>CMV LacZ</b>	7.30±4.33 (n=9)	4.49±2.89 (n=6)	5.19 ± 5.19 (n=8)	3.42 (n=1)	0.87 (n=1)
<b>CMV EGFP</b>	6.43 ±2.42 (n=2)	3.39 ±2.42 (n=2)	5.55 ±6.49 (n=4)	2.98 ±2.66 (n=2)	3.74 ±3.88 (n=2)
<b>TBG LacZ</b>	4.18 (n=1)	0.23 (n=1)	0.704±0.43 (n=2)	2.16 (n=1)	0.532 (n=1)
<b>Alb A1AT</b>	4.67±0.75 (n=2)	4.77 (n=1)	4.09 (n=1)	5.04 (n=1)	2.02 (n=1)
<b>CB A1AT</b>	0.567 (n=1)	0.438 (n=1)	2.82 (n=1)	2.78 (n=1)	0.816 ± 0.679 (n=2)
<b>TBG rhCG</b>	8.51±6.65 (n=6)	3.47±2.09 (n=5)	5.26±3.85 (n=4)	6.52±3.08 (n=4)	1.83±0.98 (n=5)
<b>TBG cFIX</b>	1.24±1.29 (n=3)	0.63±0.394 (n=6)	3.74±2.48 (n=7)	4.05 (n=1)	15.8±15.0 (n=5)

## Example 6 - Serologic Analysis of Pseudotyped Vectors

[0198] C57BL/6 mice were injected with vectors of different serotypes of AAVCBA1AT vectors intramuscularly (5 x 10<sup>11</sup> GC) and serum samples were collected 34 days later. To test neutralizing and cross-neutralizing activity of sera to each serotype of AAV, sera was analyzed in a transduction based neutralizing antibody assay [Gao, G. P., et al., (1996) *J Virol* **70**, 8934-43]. More specifically, the presence of neutralizing antibodies was determined by assessing the ability of serum to inhibit transduction of 84-31 cells by reporter viruses (AAVCMVEGFP) of different serotypes. Specifically, the reporter virus AAVCMVEGFP of each serotype [at multiplicity of infection (MOI) that led to a transduction of 90% of indicator cells] was pre-incubated with heat-inactivated serum from animals that received different serotypes of AAV or from naïve mice. After 1-hour incubation at 37° C, viruses were added to 84-31 cells in 96 well plates for 48 or 72- hour, depending on the virus serotype. Expression of GFP was measured by Fluorolmagin (Molecular Dynamics) and quantified by Image Quant Software. Neutralizing antibody titers were reported as the highest serum dilution that inhibited transduction to less than 50%.

[0199] The availability of GFP expressing vectors simplified the development of an assay for neutralizing antibodies that was based on inhibition of transduction in a permissive cell line (i.e., 293 cells stably expressing E4 from Ad5). Sera to selected AAV serotypes were generated by intramuscular injection of the recombinant viruses. Neutralization

of AAV transduction by 1:20 and 1:80 dilutions of the antisera was evaluated (See Table 5 below). Antisera to AAV1, AAV2, AAV5 and AAV8 neutralized transduction of the serotype to which the antiserum was generated (AAV5 and AAV8 to a lesser extent than AAV1 and AAV2) but not to the other serotype (i.e., there was no evidence of cross neutralization suggesting that AAV 8 is a truly unique serotype).

Table 5. Serological Analysis of New AAV Serotypes.

		% Infection on 84-31 cells with AAVCMVEGFP virus:									
		AAV2/1		AAV2/2		AAV2/5		AAV2/7		AAV2/8	
		Serum dilution:		Serum dilution:		Serum dilution:		Serum dilution:		Serum dilution:	
Sera:	Immunization Vector	1/20	1/80	1/20	1/80	1/20	1/80	1/20	1/80	1/20	1/80
Group 1	AAV2/1	0	0	100	100	100	100	100	100	100	100
Group 2	AAV2/2	100	100	0	0	100	100	100	100	100	100
Group 3	AAV2/5	100	100	100	100	16.5	16.5	100	100	100	100
Group 4	AAV2/7	100	100	100	100	100	100	61.5	100	100	100
Group 5	AAV2/8	100	100	100	100	100	100	100	100	26.3	60

**[0200]** Human sera from 52 normal subjects were screened for neutralization against selected serotypes. No serum sample was found to neutralize AAV2/7 and AAV2/8 while AAV2/2 and AAV2/1 vectors were neutralized in 20% and 10% of sera, respectively. A fraction of human pooled IgG representing a collection of 60,000 individual samples did not neutralize AAV2/7 and AAV2/8, whereas AAV2/2 and AAV2/1 vectors were neutralized at titers of serum equal to 1/1280 and 1/640, respectively.

#### Example 7 - *In vivo* Evaluation of Different Serotypes of AAV Vectors

**[0201]** In this study, 7 recombinant AAV genomes, AAV2CBhA1AT, AAV2A1bhA1AT, AAV2CMVrhCG, AAV2TBGrhCG, AAV2TBGcFIX, AAV2CMVLacZ and AAV2TBGLacZ were packaged with capsid proteins of different serotypes. In all 7 constructs, minigene cassettes were flanked with AAV2 ITRs. cDNAs of human  $\alpha$ -antitrypsin (A1AT) [Xiao, W., et al., (1999) *J Virol* **73**, 3994-4003]  $\beta$ -subunit of rhesus monkey choriogonadotropic hormone (CG) [Zoltick, P. W. & Wilson, J. M. (2000) *Mol Ther* **2**, 657-9] canine factor IX [Wang, L., et al., (1997) *Proc Natl Acad Sci USA* **94**, 11563-6] and bacterial  $\beta$ -galactosidase (i.e., Lac Z) genes were used as reporter genes. For liver-directed gene transfer, either mouse albumin gene promoter (Alb) [Xiao, W. (1999), cited above] or human thyroid hormone binding globulin gene promoter (TBG) [Wang (1997), cited above] was used to drive liver specific expression of reporter genes. In muscle-directed gene transfer experiments, either cytomegalovirus early promoter (CMV) or chicken  $\beta$ -actin promoter with CMV enhancer (CB) was employed to direct expression of reporters.

**[0202]** For muscle-directed gene transfer, vectors were injected into the right tibialis anterior of 4-6 week old NCR nude or C57BL/6 mice (Taconic, Germantown, NY). In liver-directed gene transfer studies, vectors were infused intraportally into 7-9 week old NCR nude or C57BL/6 mice (Taconic, Germantown, NY). Serum samples were collected intraorbitally at different time points after vector administration. Muscle and liver tissues were harvested at different time points for cryosectioning and Xgal histochemical staining from animals that received the lacZ vectors. For the re-administration experiment, C56BL/6 mice initially received AAV2/1, 2/2, 2/5, 2/7 and 2/8CBA1AT vectors intramuscularly and followed for A1AT gene expression for 7 weeks. Animals were then treated with AAV2/8TBGcFIX intraportally and studied for cFIX gene expression.

**[0203]** ELISA based assays were performed to quantify serum levels of hA1AT, rhCG and cFIX proteins as described previously [Gao, G. P., et al., (1996) *J Virol* **70**, 8934-43; Zoltick, P. W. & Wilson, J. M. (2000) *Mol Ther* **2**, 657-9; Wang, L., et al., *Proc Natl Acad Sci USA* **94**, 11563-6]. The experiments were completed when animals were sacrificed for harvest of muscle and liver tissues for DNA extraction and quantitative analysis of genome copies of vectors present in target tissues by TaqMan using the same set of primers and probe as in titration of vector preparations [Zhang, Y., et al., (2001) *Mol Ther* **3**, 697-707].

**[0204]** The performance of vectors based on the new serotypes were evaluated in murine models of muscle and liver-directed gene transfer and compared to vectors based on the known serotypes AAV1, AAV2 and AAV5. Vectors expressing secreted proteins ( $\alpha$ -antitrypsin (A1AT) and chorionic gonadotropin (CG)) were used to quantitate relative transduction efficiencies between different serotypes through ELISA analysis of sera. The cellular distribution of transduction within the target organ was evaluated using lacZ expressing vectors and X-gal histochemistry.

**[0205]** The performance of AAV vectors in skeletal muscle was analyzed following direct injection into the tibialis anterior muscles. Vectors contained the same AAV2 based genome with the immediate early gene of CMV or a CMV

enhanced  $\beta$ -actin promoter driving expression of the transgene. Previous studies indicated that immune competent C57BL/6 mice elicit limited humoral responses to the human A1AT protein when expressed from AAV vectors [Xiao, W., et al., (1999) *J Virol* **73**, 3994-4003].

[0206] In each strain, AAV2/1 vector produced the highest levels of A1AT and AAV2/2 vector the lowest, with AAV2/7 and AAV2/8 vectors showing intermediate levels of expression. Peak levels of CG at 28 days following injection of nu/nu NCR mice showed the highest levels from AAV2/7 and the lowest from AAV2/2 with AAV2/8 and AAV2/1 in between. Injection of AAV2/1 and AAV2/7 lacZ vectors yielded gene expression at the injection sites in all muscle fibers with substantially fewer lacZ positive fibers observed with AAV2/2 and AAV 2/8 vectors. These data indicate that the efficiency of transduction with AAV2/7 vectors in skeletal muscle is similar to that obtained with AAV2/1, which is the most efficient in skeletal muscle of the previously described serotypes [Xiao, W. (1999), cited above; Chao, H., et al., (2001) *Mol Ther* **4**, 217-22; Chao, H., et al., (2000) *Mol Ther* **2**, 619-23].

[0207] Similar murine models were used to evaluate liver-directed gene transfer. Identical doses of vector based on genome copies were infused into the portal veins of mice that were analyzed subsequently for expression of the transgene. Each vector contained an AAV2 based genome using previously described liver-specific promoters (i.e., albumin or thyroid hormone binding globulin) to drive expression of the transgene. More particularly, CMVCG and TBGCG minigene cassettes were used for muscle and liver-directed gene transfer, respectively. Levels of rhCG were defined as relative units (RUs  $\times 10^3$ ). The data were from assaying serum samples collected at day 28, post vector administration (4 animals per group). As shown in Table 3, the impact of capsid proteins on the efficiency of transduction of A1AT vectors in nu/nu and C57BL/6 mice and CG vectors in C57BL/6 mice was consistent (See Table 6).

Table 6.

Expression of $\beta$ -unit of Rhesus Monkey Chorionic Gonadotropin (rhCG)		
Vector	Muscle	Liver
AAV2/1	$4.5 \pm 2.1$	$1.6 \pm 1.0$
AAV2	$0.5 \pm 0.1$	$0.7 \pm 0.3$
AAV2/5	ND*	$4.8 \pm 0.8$
AAV2/7	$14.2 \pm 2.4$	$8.2 \pm 4.3$
AAV2/8	$4.0 \pm 0.7$	$76.0 \pm 22.8$

\* Not determined in this experiment.

[0208] In all cases, AAV2/8 vectors yielded the highest levels of transgene expression that ranged from 16 to 110 greater than what was obtained with AAV2/2 vectors; expression from AAV2/5 and AAV2/7 vectors was intermediate with AAV2/7 higher than AAV2/5. Analysis of X-Gal stained liver sections of animals that received the corresponding lacZ vectors showed a correlation between the number of transduced cells and overall levels of transgene expression. DNAs extracted from livers of C57BL/6 mice who received the A1AT vectors were analyzed for abundance of vector DNA using real time PCR technology.

[0209] The amount of vector DNA found in liver 56 days after injection correlated with the levels of transgene expression (See Table 7). For this experiment, a set of probe and primers targeting the SV40 polyA region of the vector genome was used for TaqMan PCR. Values shown are means of three individual animals with standard deviations. The animals were sacrificed at day 56 to harvest liver tissues for DNA extraction. These studies indicate that AAV8 is the most efficient vector for liver-directed gene transfer due to increased numbers of transduced hepatocytes.

Table 7 -

Real Time PCR Analysis for Abundance of AAV Vectors in nu/nu Mouse Liver Following Injection of $1 \times 10^{11}$ Genome Copies of Vector.	
AAV vectors/Dose	Genome Copies per Cell
AAV2/1A1bA1AT	$0.6 \pm 0.36$
AAV2A1bA1AT	$0.003 \pm 0.001$
AAV2/5A1bA1AT	$0.83 \pm 0.64$
AAV2/7A1bA1AT	$2.2 \pm 1.7$
AAV2/8A1bA1AT	$18 \pm 11$

[0210] The serologic data described above suggest that AAV2/8 vector should not be neutralized *in vivo* following immunization with the other serotypes. C57BL/6 mice received intraportal injections of AAV2/8 vector expressing canine

factor IX ( $10^{11}$  genome copies) 56 days after they received intramuscular injections of A1AT vectors of different serotypes. High levels of factor IX expression were obtained 14 days following infusion of AAV2/8 into naive animals ( $17 \pm 2$   $\mu\text{g/ml}$ ,  $n=4$ ) which were not significantly different that what was observed in animals immunized with AAV2/1 ( $31 \pm 23$   $\mu\text{g/ml}$ ,  $n=4$ ), AAV2/2 ( $16$   $\mu\text{g/ml}$ ,  $n=2$ ), and AAV2/7 ( $12$   $\mu\text{g/ml}$ ,  $n=2$ ). This contrasts to what was observed in AAV2/8 immunized animals that were infused with the AAV2/8 factor IX vector in which no detectable factor IX was observed ( $< 0.1$   $\mu\text{g/ml}$ ,  $n=4$ ).

**[0211]** Oligonucleotides to conserved regions of the cap gene did amplify sequences from rhesus monkeys that represented unique AAVs. Identical cap signature sequences were found in multiple tissues from rhesus monkeys derived from at least two different colonies. Full-length rep and cap open reading frames were isolated and sequenced from single sources. Only the cap open reading frames of the novel AAVs were necessary to evaluate their potential as vectors because vectors with the AAV7 or AAV8 capsids were generated using the ITRs and rep from AAV2. This also simplified the comparison of different vectors since the actual vector genome is identical between different vector serotypes. In fact, the yields of recombinant vectors generated using this approach did not differ between serotypes.

**[0212]** Vectors based on AAV7 and AAV8 appear to be immunologically distinct (i.e., they are not neutralized by antibodies generated against other serotypes). Furthermore, sera from humans do not neutralize transduction by AAV7 and AAV8 vectors, which is a substantial advantage over the human derived AAVs currently under development for which a significant proportion of the human population has pre-existing immunity that is neutralizing [Chirmule, N., et al., (1999) *Gene Ther* 6, 1574-83].

**[0213]** The tropism of each new vector is favorable for *in vivo* applications. AAV2/7 vectors appear to transduce skeletal muscle as efficiently as AAV2/1, which is the serotype that confers the highest level of transduction in skeletal muscle of the primate AAVs tested to date [Xiao, W., cited above; Chou (2001), cited above, and Chou (2000), cited above]. Importantly, AAV2/8 provides a substantial advantage over the other serotypes in terms of efficiency of gene transfer to liver that until now has been relatively disappointing in terms of the numbers of hepatocytes stably transduced. AAV2/8 consistently achieved a 10 to 100-fold improvement in gene transfer efficiency as compared to the other vectors. The basis for the improved efficiency of AAV2/8 is unclear, although it presumably is due to uptake via a different receptor that is more active on the basolateral surface of hepatocytes. This improved efficiency will be quite useful in the development of liver-directed gene transfer where the number of transduced cells is critical, such as in urea cycle disorders and familial hypercholesterolemia.

**[0214]** Thus, the present invention provides a novel approach for isolating new AAVs based on PCR retrieval of genomic sequences. The amplified sequences were easily incorporated into vectors and tested in animals. The lack of pre-existing immunity to AAV7 and the favorable tropism of the vectors for muscle indicates that AAV7 is suitable for use as a vector in human gene therapy and other *in vivo* applications. Similarly, the lack of pre-existing immunity to the AAV serotypes of the invention, and their tropisms, renders them useful in delivery of therapeutic molecules and other useful molecules.

#### Example 9 - Tissue Tropism Studies

**[0215]** In the design of a high throughput functional screening scheme for novel AAV constructs, a non-tissue specific and highly active promoter, CB promoter (CMV enhanced chicken  $\beta$  actin promoter) was selected to drive an easily detectable and quantifiable reporter gene, human  $\alpha$  anti-trypsin gene. Thus only one vector for each new AAV clone needs to be made for gene transfer studies targeting 3 different tissues, liver, lung and muscle to screen for tissue tropism of a particular AAV construct. The following table summarizes data generated from 4 novel AAV vectors in the tissue tropism studies (AAVCBA1AT), from which a novel AAV capsid clone, 44.2, was found to be a very potent gene transfer vehicle in all 3 tissues with a big lead in the lung tissue particularly. Table 8 reports data obtained (in  $\mu\text{g}$  A1AT/mL serum) at day 14 of the study.

Table 8

Vector	Target Tissue		
	Lung	Liver	Muscle
AAV2/1	ND	ND	$45 \pm 11$
AAV2/5	$0.6 \pm 0.2$	ND	ND
AAV2/8	ND	$84 \pm 30$	ND
AAV2/rh.2 (43.1)	$14 \pm 7$	$25 \pm 7.4$	$35 \pm 14$
AAV2/rh.10 (44.2)	$23 \pm 6$	$53 \pm 19$	$46 \pm 11$

## EP 1 310 571 A2

Table 8 (continued)

Vector	Target Tissue		
	Lung	Liver	Muscle
AAV2/rh.13 (42.2)	3.5±2	2±0.8	3.5±1.7
AAV2/rh.21 (42.10)	3.1±2	2±1.4	4.3±2

A couple of other experiments were then performed to confirm the superior tropism of AAV 44.2 in lung tissue. First, AAV vector carried CC10hA1AT minigene for lung specific expression were pseudotyped with capsids of novel AAVs were given to Immune deficient animals (NCR nude) in equal volume (50 µl each of the original preps without dilution) via intratracheal injections as provided in the following table. In Table 9, 50 µl of each original prep per mouse, NCR Nude, detection limit ≥0.033 µg/ml, Day 28

Table 9

Vector	Total GC in 50 µl vector	µg of A1AT/ml with 50µl vector	µg of A1AT/ml with 1x10 <sup>11</sup> vector	Relative Gene transfer as compared to rh.10 (clone 44.2)
2/1	3x10 <sup>12</sup>	2.6±0.5	0.09±0.02	2.2
2/2	5.5x10 <sup>11</sup>	<0.03	<0.005	<0.1
2/5	3.6x10 <sup>12</sup>	0.65±0.16	0.02±0.004	0.5
2/7	4.2x10 <sup>12</sup>	1±0.53	0.02±0.01	0.5
2/8	7.5x10 <sup>11</sup>	0.9±0.7	0.12±0.09	2.9
2/ch.5 (A.3.1)	9x10 <sup>12</sup>	1±0.7	0.01±0.008	0.24
2/rh.8 (43.25)	4.6x10 <sup>12</sup>	26±21	0.56±0.46	13.7
2/rh.10 (44.2)	2.8x10 <sup>12</sup>	115±38	4.1±1.4	100
2/rh.13 (42.2)	6x10 <sup>12</sup>	7.3±0.8	0.12±0.01	2.9
2/rh.21 (42.10)	2.4x10 <sup>12</sup>	9±0.9	0.38±0.04	9.3
2/rh.22 (42.11)	2.6x10 <sup>12</sup>	6±0.4	0.23±0.02	5.6
2/rh.24 (42.13)	1.1x10 <sup>11</sup>	0.4±0.3	0.4±0.3	1

The vectors were also administered to immune competent animals (C57BL/6) in equal genome copies (1x10<sup>11</sup> GC) as shown in the Table 10. (1x10<sup>11</sup> GC per animal, C57BL/6, day 14, detection limit ≥0.033 µg/ml)

Table 10

AAV Vector	µg of A1AT/ml with 1x10 <sup>11</sup> vector	Relative Gene transfer as compared to rh.10 (clone 44.2)
2/1	0.076±0.031	2.6
2/2	0.1±0.09	3.4
2/5	0.084±0.033	2.9
2/7	0.33±0.01	11
2/8	1.92±1.3	2.9
2/ch.5 (A.3.1)	0.048±0.004	1.6
2/rh.8 (43.25)	1.7±0.7	58
2/rh.10 (44.2)	2.93±1.7	100
2/rh.13 (42.2)	0.45±0.15	15
2/rh.21 (42.10)	0.86±0.32	29

# EP 1 310 571 A2

Table 10 (continued)

AAV Vector	$\mu\text{g of A1AT/ml with } 1 \times 10^{11} \text{ vector}$	Relative Gene transfer as compared to rh.10 (clone 44.2)
2/rh.22 (42.11)	$0.38 \pm 0.18$	13
2/rh.24 (42.13)	$0.3 \pm 0.19$	10

[0216] The data from both experiments confirmed the superb tropism of clone 44.2 in lung-directed gene transfer.

[0217] Interestingly, performance of clone 44.2 in liver and muscle directed gene transfer was also outstanding, close to that of the best liver transducer, AAV8 and the best muscle transducer AAV1, suggesting that this novel AAV has some intriguing biological significance.

[0218] To study serological properties of those novel AAVs, pseudotyped AAVGFP vectors were created for immunization of rabbits and in vitro transduction of 84-31 cells in the presence and absence of antisera against different capsids. The data are summarized below:

Table 11a.

Cross-NAB assay in 8431 cells and adenovirus (Adv) coinfection Infection in 8431 cells (coinfecting with Adv) with:				
Serum from rabbit immunized with:	$10^9$ GC	$10^9$ GC	$10^9$ GC	$10^{10}$ GC
	<b>rh.13</b>	<b>rh.21</b>	<b>rh.22</b>	<b>rh.24</b>
	AAV2/42.2	AAV2/42.10	AAV2/42.11	AAV2/42.13
AAV2/1	1/20	1/20	1/20	No NAB
AAV2/2	1/640	1/1280	1/5120	No NAB
AAV2/5	No NAB	1/40	1/160	No NAB
AAV2/7	1/81920	1/81920	1/40960	1/640
AAV2/8	1/640	1/640	1/320	1/5120
<b>Ch.5 AAV2/A3</b>	1/20	1/160	1/640	1/640
<b>rh.8</b> AAV2/43.25	1/20	1/20	1/20	1/320
<b>rh.10</b> AAV2/44.2	No NAB	No NAB	No NAB	1/5120
<b>rh.13</b> AAV2/42.2	1/5120	1/5120	1/5120	No NAB
<b>rh.21</b> AAV2/42.10	1/5120	1/10240	1/5120	1/20
<b>rh.22</b> AAV2/42.11	1/20480	1/20480	1/40960	No NAB
<b>rh.24</b> AAV2/42.13	No NAB	1/20	1/20	1/5120

EP 1 310 571 A2

Table 11b.

Cross-NAB assay in 8431 cells and Adv coinfection Infection in 8431 cells (coinfecting with Adv) with:					
Serum from rabbit immunized with:	10 <sup>9</sup> GC	10 <sup>10</sup> GC	10 <sup>10</sup> GC	10 <sup>9</sup> GC	10 <sup>9</sup> GC
	<b>rh.12</b>	<b>ch.5</b>	<b>rh.8</b>	<b>rh.10</b>	<b>rh.20</b>
	AAV2/42.1B	AAV2/A3	AAV2/43.25	AAV2/44.2	AAV2/42.8.2
AAV2/1	No NAB	1/20480	No NAB	1/80	ND
AAV2/2	1/20	No NAB	No NAB	No NAB	ND
AAV2/5	No NAB	1/320	No NAB	No NAB	ND
AAV2/7	1/2560	1/640	1/160	1/81920	ND
AAV2/8	1/10240	1/2560	1/2560	1/81920	ND
<b>ch.5</b> AAV2/A3	1/1280	1/10240	ND	1/5120	1/320
<b>rh.8</b> AAV2/43.25	1/1280	ND	1/20400	1/5120	1/2560
<b>rh.10</b> AAV2/44.2	1/5120	ND	ND	1/5120	1/5120
<b>rh.13</b> AAV2/42.2	1/20	ND	ND	No NAB	1/320
<b>rh.21</b> AAV2/42.10	1/20	ND	ND	1/40	1/80
<b>rh.22</b> AAV2/42.11	No NAB	ND	ND	ND	No NAB
<b>rh.24</b> AAV2/42.13	1/5120	ND	ND	ND	1/2560

Table 12

Titer of rabbit sera			Titer after Boosting
Vector		Titer d21	
<b>ch.5</b>	AAV2/A3	1/10,240	1/40,960
<b>rh.8</b>	AAV2/43.25	1/20,400	1/163,840
<b>rh.10</b>	AAV2/44.2	1/10,240	1/527,680
<b>rh.13</b>	AAV2/42.2	1/5,120	1/20,960
<b>rh.21</b>	AAV2/42.10	1/20,400	1/81,920
<b>rh.22</b>	AAV2/42.11	1/40,960	ND
<b>rh.24</b>	AAV2/42.13	1/5,120	ND

Table 13 a.

Infection in 8431 cells (coinfecting with Adv) with GFP						
	10 <sup>9</sup> GC/well	10 <sup>9</sup> GC/well	10 <sup>9</sup> GC/well	10 <sup>9</sup> GC/well	10 <sup>9</sup> GC/well	10 <sup>9</sup> GC/well
						<b>ch.5</b>
	AAV2/1	AAV2/2	AAV2/5	AAV2/7	AAV2/8	AAV2/A3
# GFU/field	128	>200	95	56	13	1
	83	>200	65	54	11	1

Table 13b.

Infection in 8431 cells (coinfecting with Adv) with GFP							
	10 <sup>9</sup> GC/well	10 <sup>9</sup> GC/ well	10 <sup>9</sup> GC/ well	10 <sup>9</sup> GC/well	10 <sup>9</sup> GC/well	10 <sup>9</sup> GC/well	10 <sup>9</sup> GC/well
	<i>rh.8</i>	<i>rh.10</i>	<i>rh.13</i>	<i>rh.21</i>	<i>rh.22</i>	<i>rh.24</i>	<i>rh.12</i>
	AAV2/43.25	AAV2/44.2	AAV2/42.2	AAV2/42.10	AAV2/42.11	AAV2/42.13	AAV2/42.1B
#	3	13	54	62	10	3	18
GFU/ field	2	12	71	60	14	2	20
			48	47	16	3	12

#### Example 10 - Mouse Model of Familial Hypercholesterolemia

**[0219]** The following experiment demonstrates that the AAV2/7 construct of the invention delivers the LDL receptor and express LDL receptor in an amount sufficient to reduce the levels of plasma cholesterol and triglycerides in animal models of familial hypercholesterolemia.

##### A. Vector Construction

**[0220]** AAV vectors packaged with AAV7 or AAV8 capsid proteins were constructed using a pseudotyping strategy [Hildinger M, *et al.*, *J. Virol* 2001; **75**:6199-6203]. Recombinant AAV genomes with AAV2 inverted terminal repeats (ITR) were packaged by triple transfection of 293 cells with the *cis*-plasmid, the adenovirus helper plasmid and a chimeric packaging construct, a fusion of the capsids of the novel AAV serotypes with the rep gene of AAV2. The chimeric packaging plasmid was constructed as previously described [Hildinger et al, cited above]. The recombinant vectors were purified by the standard CsCl<sub>2</sub> sedimentation method. To determine the yield TaqMan (Applied Biosystems) analysis was performed using probes and primers targeting the SV40 poly(A) region of the vectors [Gao GP, *et al.*, *Hum Gene Ther.* 2000 Oct 10;**11**(15):2079-91]. The resulting vectors express the transgene under the control of the human thyroid hormone binding globulin gene promoter (TBG).

##### B. Animals

**[0221]** LDL receptor deficient mice on the C57B1/6 background were purchased from the Jackson Laboratory (Bar Harbor, ME, USA) and maintained as a breeding colony. Mice were given unrestricted access to water and obtained a high fat Western Diet (high % cholesterol) starting three weeks prior vector injection. At day -7 as well at day 0, blood was obtained via retroorbital bleeds and the lipid profile evaluated. The mice were randomly divided into seven groups. The vector was injected via an intraportal injection as previously described ([Chen SJ *et al.*, *Mol Therapy* 2000; **2**(3), 256-261]. Briefly, the mice were anaesthetized with ketamine and xylazine. A laparotomy was performed and the portal vein exposed. Using a 30g needle the appropriate dose of vector diluted in 100ul PBS was directly injected into the portal vein. Pressure was applied to the injection site to ensure a stop of the bleeding. The skin wound was closed and draped and the mice carefully monitored for the following day. Weekly bleeds were performed starting at day 14 after liver directed gene transfer to measure blood lipids. Two animals of each group were sacrificed at the time points week 6 and week 12 after vector injection to examine atherosclerotic plaque size as well as receptor expression. The remaining mice were sacrificed at week 20 for plaque measurement and determination of transgene expression.

Table 14

	Vector	dose	n
Group 1	AAV2/7-TBG-hLDLr	1x 10 <sup>12</sup> gc	12
Group 2	AAV2/7-TBG-hLDLr	3x 10 <sup>11</sup> gc	12
Group 3	AAV2/7-TBG-hLDLr	1x 10 <sup>11</sup> gc	12
Group 4	AAV2/8-TBG-hLDLr	1x 10 <sup>12</sup> gc	12
Group 5	AAV2/8-TBG-hLDLr	3x 10 <sup>11</sup> gc	12

Table 14 (continued)

	<i>Vector</i>	<i>dose</i>	<i>n</i>
Group 6	AAV2/8-TBG-hLDLr	1x 10 <sup>11</sup> gc	12
Group 7	AAV2/7-TBG-LacZ	1x 10 <sup>11</sup> gc	16

### C. Serum lipoprotein and liver function analysis

**[0222]** Blood samples were obtained from the retroorbital plexus after a 6 hour fasting period. The serum was separated from the plasma by centrifugation. The amount of plasma lipoproteins and liver transaminases in the serum were detected using an automatized clinical chemistry analyzer (ACE, Schiapparelli Biosystems, Alpha Wassermann)

### D. Detection of transgene expression

**[0223]** LDL receptor expression was evaluated by immuno-fluorescence staining and Western blotting. For Western Blot frozen liver tissue was homogenized with lysis buffer ( 20 mM Tris, pH7.4, 130mM NaCl, 1% Triton X 100, protease inhibitor (complete, EDTA-free, Roche, Mannheim, Germany). Protein concentration was determined using the Micro BCA Protein Assay Reagent Kit (Pierce, Rockford, IL). 40 µg of protein was resolved on 4- 15% Tris-HCl Ready Gels (Biorad, Hercules, CA) and transferred to a nitrocellulose membrane (Invitrogen, ). To generate Anti-hLDL receptor antibodies a rabbit was injected intravenously with an AdhLDLr prep (1x10<sup>13</sup> GC). Four weeks later the rabbit serum was obtained and used for Western Blot. A 1:100 dilution of the serum was used as a primary antibody followed by a HRP-conjugated anti-rabbit IgG and ECL chemiluminescent detection (ECL Western Blot Detection Kit, Amersham, Arlington Heights, IL).

### E. Immunocytochemistry

**[0224]** For determination of LDL receptor expression in frozen liver sections immunohistochemistry analyses were performed. 10µm cryostat sections were either fixed in acetone for 5 minutes, or unfixed. Blocking was obtained *via* a 1 hour incubation period with 10% of goat serum. Sections were then incubated for one hour with the primary antibody at room temperature. A rabbit polyclonal antibody anti-human LDL (Biomedical Technologies Inc., Stoughton, MA) was used diluted accordingly to the instructions of the manufacturer. The sections were washed with PBS, and incubated with 1:100 diluted fluorescein goat anti-rabbit IgG (Sigma, St Louis, MO). Specimens were finally examined under fluorescence microscope Nikon Microphot-FXA. In all cases, each incubation was followed by extensive washing with PBS. Negative controls consisted of preincubation with PBS, omission of the primary antibody, and substitution of the primary antibody by an isotype-matched non-immune control antibody. The three types of controls mentioned above were performed for each experiment on the same day.

### F. Gene transfer efficiency

**[0225]** Liver tissue was obtained after sacrificing the mice at the designated time points. The tissue was shock frozen in liquid nitrogen and stored at -80°C until further processing. DNA was extracted from the liver tissue using a QIAamp DNA Mini Kit (QIAGEN GmbH, Germany) according to the manufacturers protocol. Genome copies of AAV vectors in the liver tissue were evaluated using Taqman analysis using probes and primers against the SV40 poly(A) tail as described above.

### G. Atherosclerotic plaque measurement

**[0226]** For the quantification of the atherosclerotic plaques in the mouse aorta the mice were anaesthetized (10% ketamine and xylazine, ip), the chest opened and the arterial system perfused with ice-cold phosphate buffered saline through the left ventricle. The aorta was then carefully harvested, slit down along the ventral midline from the aortic arch down to the femoral arteries and fixed in formalin. The lipid-rich atherosclerotic plaques were stained with Sudan IV (Sigma, Germany) and the aorta pinned out flat on a black wax surface. The image was captured with a Sony DXC-960 MD color video camera. The area of the plaque as well as of the complete aortic surface was determined using Phase 3 Imaging Systems (Media Cybernetics).

H. Clearance of  $I^{125}$  LDL

[0227] Two animals per experimental group were tested. A bolus of  $I^{125}$  - labeled LDL (generously provided by Dan Rader, U Penn) was infused slowly through the tail vein over a period of 30 sec (1,000,000 counts of  $[I^{125}]$ -LDL diluted in 100 $\mu$ l sterile PBS/ animal). At time points 3min, 30 min, 1.5hr, 3hr, 6hr after injection a blood sample was obtained via the retro-orbital plexus. The plasma was separated off from the whole blood and 10 $\mu$ l plasma counted in the gamma counter. Finally the fractional catabolic rate was calculated from the lipoprotein clearance data.

## I. Evaluation of Liver Lipid accumulation

[0228] Oil Red Staining of frozen liver sections was performed to determine lipid accumulation. The frozen liver sections were briefly rinsed in distilled water followed by a 2 minute incubation in absolute propylene glycol. The sections were then stained in oil red solution (0.5% in propylene glycol) for 16 hours followed by counterstaining with Mayer's hematoxylin solution for 30 seconds and mounting in warmed glycerin jelly solution.

[0229] For quantification of the liver cholesterol and triglyceride content liver sections were homogenized and incubated in chloroform/methanol (2:1) overnight. After adding of 0.05%  $H_2SO_4$  and centrifugation for 10 minutes, the lower layer of each sample was collected, divided in two aliquots and dried under nitrogen. For the cholesterol measurement the dried lipids of the first aliquot were dissolved in 1% Triton X-100 in chloroform. Once dissolved, the solution was dried under nitrogen. After dissolving the lipids in ddH<sub>2</sub>O and incubation for 30 minutes at 37°C the total cholesterol concentration was measured using a Total Cholesterol Kit (Wako Diagnostics). For the second aliquot the dried lipids were dissolved in alcoholic KOH and incubated at 60°C for 30 minutes. Then 1M MgCl<sub>2</sub> was added, followed by incubation on ice for 10 minutes and centrifugation at 14,000 rpm for 30 minutes. The supernatant was finally evaluated for triglycerides (Wako Diagnostics).

[0230] All of the vectors pseudotyped in an AAV2/8 or AAV2/7 capsid lowered total cholesterol, LDL and triglycerides as compared to the control. These test vectors also corrected phenotype of hypercholesterolemia in a dose-dependent manner. A reduction in plaque area for the AAV2/8 and AAV2/7 mice was observed in treated mice at the first test (2 months), and the effect was observed to persist over the length of the experiment (6 months).

## Example 10 - Functional Factor IX Expression and Correction of Hemophilia

## A. Knock-Out Mice

[0231] Functional canine factor IX (FIX) expression was assessed in hemophilia B mice. Vectors with capsids of AAV1, AAV2, AAV5, AAV7 or AAV8 were constructed to deliver AAV2 5' ITR - liver-specific promoter [LSP] - canine FIX - woodchuck hepatitis post-regulatory element (WPRE) - AAV2 3' ITR. The vectors were constructed as described in Wang et al, 2000, *Molecular Therapy* 2: 154-158), using the appropriate capsids.

[0232] Knock-out mice were generated as described in Wang et al, 1997. *Proc. Natl. Acad. Sci. USA* 94: 11563-11566. This model closely mimic the phenotypes of hemophilia B in human.

[0233] Vectors of different serotypes (AAV1, AAV2, AAV5, AAV7 and AAV8) were delivered as a single intraportal injection into the liver of adult hemophilic C57B1/6 mice in a dose of  $1 \times 10^{11}$  GC/mouse for the five different serotypes and one group received an AAV8 vector at a lower dose,  $1 \times 10^{10}$  GC/mouse. Control group was injected with  $1 \times 10^{11}$  GC of AAV2/8 TBG LacZ3. Each group contains 5-10 male and female mice. Mice were bled bi-weekly after vector administration.

## 1. ELISA

[0234] The canine FIX concentration in the mouse plasma was determined by an ELISA assay specific for canine factor IX, performed essentially as described by Axelrod et al, 1990, *Proc. Natl. Acad. Sci. USA*, 87:5173-5177 with modifications. Sheep anti-canine factor IX (Enzyme Research Laboratories) was used as primary antibody and rabbit anti-canine factor IX ((Enzyme Research Laboratories) was used as secondary antibody. Beginning at two weeks following injection, increased plasma levels of cFIX were detected for all test vectors. The increased levels were sustained at therapeutic levels throughout the length of the experiment, i.e., to 12 weeks. Therapeutic levels are considered to be 5% of normal levels, i.e., at about 250 ng/mL.

[0235] The highest levels of expression were observed for the AAV2/8 (at  $10^{11}$ ) and AAV2/7 constructs, with sustained superphysiology levels cFIX levels (ten-fold higher than the normal level). Expression levels for AAV2/8 ( $10^{11}$ ) were approximately 10 fold higher than those observed for AAV2/2 and AAV2/8 ( $10^{10}$ ). The lowest expression levels, although still above the therapeutic range, were observed for AAV2/5.

## 2. *In Vitro* Activated Partial Thromboplastin time (aPTT) Assay

**[0236]** Functional factor IX activity in plasma of the FIX knock-out mice was determined by an *in vitro* activated partial thromboplastin time (aPTT) assay-Mouse blood samples were collected from the retro-orbital plexus into 1/10 volume of citrate buffer. The aPTT assay was performed as described by Wang et al, 1997, *Proc. Natl. Acad. Sci. USA* **94**: 11563-11566.

**[0237]** Clotting times by aPTT on plasma samples of all vector injected mice were within the normal range (approximately 60 sec) when measured at two weeks post-injection, and sustained clotting times in the normal or shorter than normal range throughout the study period (12 weeks).

**[0238]** Lowest sustained clotting times were observed in the animals receiving AAV2/8 ( $10^{11}$ ) and AAV2/7. By week 12, AAV2/2 also induced clotting times similar to those for AAV2/8 and AAV2/7. However, this lowered clotting time was not observed for AAV2/2 until week 12, whereas lowered clotting times (in the 25 - 40 sec range) were observed for AAV2/8 and AAV2/7 beginning at week two.

**[0239]** Immuno-histochemistry staining on the liver tissues harvested from some of the treated mice is currently being performed. About 70-80% of hepatocytes are stained positive for canine FIX in the mouse injected with AAV2/8.cFIX vector.

## B. *Hemophilia B Dogs*

**[0240]** Dogs that have a point mutation in the catalytic domain of the F.IX gene, which, based on modeling studies, appears to render the protein unstable, suffer from hemophilia B [Evans et al, 1989, *Proc. Natl. Acad. Sci. USA*, **86**: 10095-10099]. A colony of such dogs has been maintained for more than two decades at the University of North Carolina, Chapel Hill. The homeostatic parameters of these dogs are well described and include the absence of plasma F.IX antigen, whole blood clotting times in excess of 60 minutes, whereas normal dogs are 6-8 minutes, and prolonged activated partial thromboplastin time of 50-80 seconds, whereas normal dogs are 13-28 seconds. These dogs experience recurrent spontaneous hemorrhages. Typically, significant bleeding episodes are successfully managed by the single intravenous infusion of 10 ml/kg of normal canine plasma; occasionally, repeat infusions are required to control bleeding.

**[0241]** Four dogs are injected intraportally with AAV.cFIX according to the schedule below. A first dog receives a single injection with AAV2/2.cFIX at a dose of  $3.7 \times 10^{11}$  genome copies (GC)/kg. A second dog receives a first injection of AAV2/2.cFIX ( $2.8 \times 10^{11}$  GC/kg), followed by a second injection with AAV2/7.cFIX ( $2.3 \times 10^{13}$  GC/kg) at day 1180. A third dog receives a single injection with AAV2/2.cFIX at a dose of  $4.6 \times 10^{12}$  GC/kg. The fourth dog receives an injection with AAV2/2.cFIX ( $2.8 \times 10^{12}$  GC/kg) and an injection at day 995 with AAV2/7.cFIX ( $5 \times 10^{12}$  GC/kg).

**[0242]** The abdomen of hemophilia dogs are aseptically and surgically opened under general anesthesia and a single infusion of vector is administered into the portal vein. The animals are protected from hemorrhage in the peri-operative period by intravenous administration of normal canine plasma. The dog is sedated, intubated to induce general anesthesia, and the abdomen shaved and prepped. After the abdomen is opened, the spleen is moved into the operative field. The splenic vein is located and a suture is loosely placed proximal to a small distal incision in the vein. A needle is rapidly inserted into the vein, then the suture loosened and a 5 F cannula is threaded to an intravenous location near the portal vein threaded to an intravenous location near the portal vein bifurcation. After hemostasis is secured and the catheter balloon inflated, approximately 5.0 ml of vector diluted in PBS is infused into the portal vein over a 5 minute interval. The vector infusion is followed by a 5.0 ml infusion of saline. The balloon is then deflated, the cannula removed and venous hemostasis is secured. The spleen is then replaced, bleeding vessels are cauterized and the operative wound is closed. The animal is extubated having tolerated the surgical procedure well. Blood samples are analyzed as described. [Wang et al, 2000, *Molecular Therapy* **2**: 154-158]

**[0243]** Results showing correction or partial correction are anticipated for AAV2/7.

**[0244]** All publications cited in this specification are incorporated herein by reference. While the invention has been described with reference to a particularly preferred embodiment, it will be appreciated that modifications can be made without departing from the spirit of the invention. Such modifications are intended to fall within the scope of the claims.

# EP 1 310 571 A2

## SEQUENCE LISTING

5 <110> The Trustees of The University of Pennsylvania  
 <120> A Method of Detecting and/or Identifying Adeno-Associated  
 Virus (AAV) Sequences and Isolating Novel Sequences Identified  
 Thereby  
 10 <130> UPN-02735ff  
 <150> US 60/350,607  
 <151> 2001-11-13  
 <150> US 60/341,117  
 15 <151> 2001-12-17  
 <150> US 60/377,066  
 <151> 2002-05-01  
 <150> US 60/386,675  
 20 <151> 2002-06-05  
 <160> 120  
 <170> PatentIn version 3.1  
 25 <210> 1  
 <211> 4721  
 <212> DNA  
 <213> adeno-associated virus serotype 7  
 30 <400> 1  
 ttggccactc cctctatgcg cgctcgctcg ctcggtgggg cctgcggacc aaaggtccgc 60  
 agacggcaga gctctgctct gccggcccca ccgagcgagc gagcgcgcac agagggagtg 120  
 35 gccaaactcca tcactagggg taccgcgaag cgcctccac gctgccgcgt cagcgtgac 180  
 gtaaatacacg tcatagggga gtggtcctgt attagctgtc acgtgagtgc ttttgcgaca 240  
 ttttgcgaca ccacgtggcc atttgaggta tatatggccg agtgagcgag caggatctcc 300  
 40 attttgaccg cgaaatttga acgagcagca gccatgccgg gtttctacga gatcgtgac 360  
 aaggtgccga gcgacctgga cgagcacctg ccgggcattt ctgactcgtt tgtgaactgg 420  
 gtggccgaga aggaatggga gctgcccccg gattctgaca tggatctgaa tctgatcgag 480  
 45 caggcacccc tgaccgtggc cgagaagctg cagcgcgact tcctgggtcca atggcgccgc 540  
 gtgagtaagg ccccgagggc cctgttcttt gttcagttcg agaaggcgga gagctacttc 600  
 caccttcacg ttctgggtga gaccacgggg gtcaagtcca tggtgctagg ccgcttcctg 660  
 50 agtcagattc gggagaagct ggtccagacc atctaccgcg gggtcgagcc cagcgtgcc 720  
 aactggttcg cggtgaccaa gacgcgtaat ggcgcggcg gggggaacaa ggtggtggac 780  
 gagtgtctaca tccccaacta cctcctgccc aagaccagc ccgagctgca gtgggcgtgg 840  
 55 actaacatgg aggagtatat aagcgcgtgt ttgaacctgg ccgaacgcaa acggctcgtg 900

EP 1 310 571 A2

	gcgagcacc tgaccacgt cagccagacg caggagcaga acaaggagaa tctgaacccc	960
5	aattctgacg cgcccgatgat caggtcaaaa acctccgcgc gctacatgga gctgggtcggg	1020
	tggctgggtgg accggggcat cacctccgag aagcagtggg tccaggagga ccaggcctcg	1080
	tacatctcct tcaacgccgc ctccaactcg cgggtcccaga tcaaggccgc gctggacaat	1140
10	gccggcaaga tcatggcgct gaccaaatec gcgccgact acctgggtggg gccctcgctg	1200
	cccgcggaca ttaaaaccaa ccgcatctac cgcatcctgg agctgaacgg gtacgatcct	1260
	gcctacgccg gctccgtctt tctcggctgg gccagaaaa agttcgggaa gcgcaacacc	1320
15	atctgggtgt ttggggccgc caccaccggc aagaccaaca ttgcggaagc catcgcccac	1380
	gccgtgccct tctacggctg cgtcaactgg accaatgaga actttccctt caacgattgc	1440
	gtcgacaaga tgggtgatctg gtgggaggag ggcaagatga cggccaaggt cgtggagtcc	1500
20	gccaaggcca ttctcggcgg cagcaaggtg cgcggtggacc aaaagtgcaa gtcgtccgcc	1560
	cagatcgacc ccacccccgt gatcgtcacc tccaacacca acatgtgcgc cgtgattgac	1620
	gggaacagca ccaccttcga gcaccagcag ccgttgacag accggatgtt caaatttgaa	1680
25	ctcaccgcc gtctggagca cgactttggc aagggtgacga agcaggaagt caaagagttc	1740
	ttccgctggg ccagtgatca cgtgaccgag gtggcgcatg agttctacgt cagaaagggc	1800
	ggagccagca aaagaccgc ccccgatgac gcggatataa gcgagcccaa gcgggcctgc	1860
30	ccctcagtcg cggatccatc gacgtcagac gcggaaggag ctccgggtgga ctttgccgac	1920
	aggtaccaaa acaaagtgtt tcgtcacgcg ggcatgattc agatgctgtt tccctgcaaa	1980
	acgtgcgaga gaatgaatca gaatttcaac atttgcttca cacacggggt cagagactgt	2040
35	ttagagtgtt tccccggcgt gtcagaatct caaccggctg tcagaaaaaa gacgtatcgg	2100
	aaactctgcg cgattcatca tctgctgggg cgggcgccc agattgcttg ctcggcctgc	2160
	gacctggtca acgtggacct ggacgactgc gtttctgagc aataaatgac ttaaaccagg	2220
40	tatggctgcc gatggttatc ttccagattg gctcgaggac aacctctctg agggcattcg	2280
	cgagtgggtg gacctgaaac ctggagcccc gaaacccaaa gccaacacgc aaaagcagga	2340
	caacggccgg ggtctggtgc ttccctggcta caagtacctc ggacccttca acggactcga	2400
45	caagggggag cccgtcaacg cggcgagcgc agcgccctc gagcacgaca aggcctacga	2460
	ccagcagctc aaagcgggtg acaatccgta cctgcggtat aaccacgcg acgcccagtt	2520
	tcaggagcgt ctgcaagaag atacgtcatt tgggggcaac ctcgggcgag cagtcttcca	2580
50	ggccaagaag cgggttctcg aacctctcgg tctggttgag gaaggcgcta agacggctcc	2640
	tgcaagaag agaccggtag agccgtcacc tcagcgttcc cccgactcct ccacgggcat	2700
	cggcaagaaa ggccagcagc ccgccagaaa gagactcaat ttcggtcaga ctggcgactc	2760
55	agagtcatgc cccgacctc aacctctcgg agaacctcca gcagcgccct ctagtgtggg	2820

EP 1 310 571 A2

	atctggtaca gtggctgcag gcggtggcgc accaatggca gacaataacg aaggtgccga	2880
5	cggagtgggt aatgcctcag gaaattggca ttgcgattcc acatggctgg gcgacagagt	2940
	cattaccacc agcaccgaa cctggggcct gccacctac aacaaccacc tctacaagca	3000
	aatctccagt gaaactgcag gtagtaccaa cgacaacacc tacttcggct acagaccccc	3060
10	ctgggggtat tttgacttta acagattcca ctgccacttc tcaccacgtg actggcagcg	3120
	actcatcaac aacaactggg gattccggcc caagaagctg cggttcaagc tcttcaacat	3180
	ccaggtaag gaggtcacga cgaatgacgg cgttacgacc atcgctaata accttaccag	3240
15	cacgattcag gtattctcgg actcggaata ccagctgccg tacgtcctcg gctctgcgca	3300
	ccagggtgc ctgcctcgt tccggcgga cgtcttcatg attcctcagt acggctacct	3360
	gactctcaac aatggcagtc agtctgtggg acgttcctcc ttctactgcc tggagtactt	3420
20	ccccctcag atgctgagaa cgggcaacaa ctttgagttc agctacagct tcgaggacgt	3480
	gcctttccac agcagctacg cacacagcca gagcctggac cggctgatga atcccccat	3540
	cgaccagtac ttgtactacc tggccagaac acagagtaac ccaggaggca cagctggcaa	3600
25	tcgggaactg cagttttacc agggcgggccc ttcaactatg gccgaacaag ccaagaattg	3660
	gttacctgga ccttgcttcc ggcaacaaag agtctccaaa acgctggatc aaaacaacaa	3720
	cagcaacttt gcttggaactg gtgccaccaa atatcacctg aacggcagaa actcgttggt	3780
30	taatcccggc gtcgccatgg caactcaciaa ggacgacgag gaccgctttt tcccatccag	3840
	cggagtccctg atttttggaa aaactggagc aactaacaaa actacattgg aaaatgtgtt	3900
	aatgacaaat gaagaagaaa ttctgcctac taatcctgta gccacggaag aatacgggat	3960
35	agtcagcagc aacttacaag cggctaatac tgcagcccag acacaagttg tcaacaacca	4020
	gggagccctta cctggcatgg tctggcagaa ccgggacgtg tacctgcagg gtcccatctg	4080
	ggccaagatt cctcacacgg atggcaactt tcaccogtct cctttgatgg gcggctttgg	4140
40	acttaaacad ccgcctcctc agatcctgat caagaacact cccgttcccg ctaatcctcc	4200
	ggaggtgttt actcctgcc agtttgcttc gttcatcaca cagtacagca ccggacaagt	4260
	cagcgtggaa atcgagtggg agctgcagaa ggaaaacagc aagcgttggg acccgagat	4320
45	tcagtacacc tccaactttg aaaagcagac tgggtgtggac tttgccgttg acagccaggg	4380
	tgtttactct gagcctcgcc ctattggcac tcgttacctc acccgtaatc tgtaattgca	4440
	tgttaataca taaaccgggt gattcgtttc agttgaactt tgggtctcctg tgcttcttat	4500
50	cttatcgggt tccatagcaa ctggttacac attaaactgct tgggtgcgct tcacgataag	4560
	aacactgacg tcaccgcggt acccctagtg atggagtgtg ccactccctc tatgcgcgct	4620
	cgctcgctcg gtggggcctg cggaccaaag gtccgcagac ggcagagctc tgctctgccg	4680
55	gccccaccga gcgagcgagc gcgcatagag ggagtggcca a	4721

# EP 1 310 571 A2

```

5      <210> 2
      <211> 737
      <212> PRT
      <213> capsid protein of adeno-associated virus serotype 7

      <400> 2
10     Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
      1          5          10          15

      Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
      20          25          30

15     Lys Ala Asn Gln Gln Lys Gln Asp Asn Gly Arg Gly Leu Val Leu Pro
      35          40          45

      Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
20     50          55          60

      Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
      65          70          75          80

25     Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
      85          90          95

      Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
30     100         105         110

      Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
      115         120         125

35     Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Ala Lys Lys Arg
      130         135         140

      Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
40     145         150         155         160

      Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln
      165         170         175

45     Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro
      180         185         190

      Pro Ala Ala Pro Ser Ser Val Gly Ser Gly Thr Val Ala Ala Gly Gly
50     195         200         205

      Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn
      210         215         220

55     Ala Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
      225         230         235         240

```

# EP 1 310 571 A2

5 Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His  
 245 250 255  
 Leu Tyr Lys Gln Ile Ser Ser Glu Thr Ala Gly Ser Thr Asn Asp Asn  
 260 265 270  
 10 Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg  
 275 280 285  
 15 Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn  
 290 295 300  
 Asn Trp Gly Phe Arg Pro Lys Lys Leu Arg Phe Lys Leu Phe Asn Ile  
 305 310 315 320  
 20 Gln Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn  
 325 330 335  
 25 Asn Leu Thr Ser Thr Ile Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu  
 340 345 350  
 Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro  
 355 360 365  
 30 Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn  
 370 375 380  
 35 Gly Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe  
 385 390 395 400  
 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Ser  
 405 410 415  
 40 Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu  
 420 425 430  
 45 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala  
 435 440 445  
 Arg Thr Gln Ser Asn Pro Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln  
 450 455 460  
 50 Phe Tyr Gln Gly Gly Pro Ser Thr Met Ala Glu Gln Ala Lys Asn Trp  
 465 470 475 480  
 55 Leu Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp  
 485 490 495

# EP 1 310 571 A2

5 Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His  
 500 505 510  
 Leu Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr  
 515 520 525  
 10 His Lys Asp Asp Glu Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile  
 530 535 540  
 15 Phe Gly Lys Thr Gly Ala Thr Asn Lys Thr Thr Leu Glu Asn Val Leu  
 545 550 555 560  
 Met Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu  
 565 570 575  
 20 Glu Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Asn Thr Ala Ala  
 580 585 590  
 25 Gln Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp  
 595 600 605  
 Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro  
 610 615 620  
 30 His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly  
 625 630 635 640  
 35 Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro  
 645 650 655  
 Ala Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Phe Ala Ser Phe Ile  
 660 665 670  
 40 Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu  
 675 680 685  
 45 Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser  
 690 695 700  
 Asn Phe Glu Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly  
 705 710 715 720  
 50 Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn  
 725 730 735  
 Leu  
 55

# EP 1 310 571 A2

5 <210> 3  
 <211> 623  
 <212> PRT  
 <213> rep protein of adeno-associated virus serotype 7  
  
 <400> 3  
 10 Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
  
 15 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 20  
  
 Val Gln Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
  
 25 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Leu His Val Leu Val Glu  
 85 90 95  
  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 30  
  
 Arg Glu Lys Leu Val Gln Thr Ile Tyr Arg Gly Val Glu Pro Thr Leu  
 115 120 125  
  
 35 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 40  
  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Glu Tyr Ile  
 165 170 175  
  
 45 Ser Ala Cys Leu Asn Leu Ala Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Leu Asn  
 195 200 205  
 50  
  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
  
 55 Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys  
 225 230 235 240

# EP 1 310 571 A2

5 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 10 Ile Met Ala Leu Thr Lys Ser Ala Pro Asp Tyr Leu Val Gly Pro Ser  
 275 280 285  
 15 Leu Pro Ala Asp Ile Lys Thr Asn Arg Ile Tyr Arg Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Ala Tyr Ala Gly Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 20 Gln Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 25 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 30 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 35 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 40 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 45 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 50 Glu Val Lys Glu Phe Phe Arg Trp Ala Ser Asp His Val Thr Glu Val  
 465 470 475 480  
 Ala His Glu Phe Tyr Val Arg Lys Gly Gly Ala Ser Lys Arg Pro Ala  
 485 490 495  
 55

# EP 1 310 571 A2

5 Pro Asp Asp Ala Asp Ile Ser Glu Pro Lys Arg Ala Cys Pro Ser Val  
 500 505 510  
 Ala Asp Pro Ser Thr Ser Asp Ala Glu Gly Ala Pro Val Asp Phe Ala  
 515 520 525  
 10 Asp Arg Tyr Gln Asn Lys Cys Ser Arg His Ala Gly Met Ile Gln Met  
 530 535 540  
 Leu Phe Pro Cys Lys Thr Cys Glu Arg Met Asn Gln Asn Phe Asn Ile  
 545 550 555 560  
 15 Cys Phe Thr His Gly Val Arg Asp Cys Leu Glu Cys Phe Pro Gly Val  
 565 570 575  
 20 Ser Glu Ser Gln Pro Val Val Arg Lys Lys Thr Tyr Arg Lys Leu Cys  
 580 585 590  
 Ala Ile His His Leu Leu Gly Arg Ala Pro Glu Ile Ala Cys Ser Ala  
 595 600 605  
 25 Cys Asp Leu Val Asn Val Asp Leu Asp Asp Cys Val Ser Glu Gln  
 610 615 620  
 30 <210> 4  
 <211> 4393  
 <212> DNA  
 <213> adeno-associated virus serotype 8  
 <400> 4  
 35 cagagagggga gtggccaact ccatacactag gggtagcgcg aagcgccctcc cacgctgccg 60  
 cgtcagcgct gacgtaaatt acgtcatagg ggagtgggtcc tgtattagct gtcacgtgag 120  
 tgctttttgcg gcatttttgcg acaccacgtg gccatttgag gtatatatgg ccgagtgagc 180  
 40 gagcaggatc tccattttga ccgcgaaatt tgaacgagca gcagccatgc cgggcttcta 240  
 cgagatcgtg atcaagggtgc cgagcgacct ggacgagcac ctgccgggca tttctgactc 300  
 gtttgtgaac tgggtggccg agaaggaatg ggagctgccc ccggattctg acatggatcg 360  
 45 gaatctgata gagcaggcac cctgacctg gcccgagaag ctgcagcgcg acttcttggt 420  
 ccaatggcgc cgcgtgagta aggcccgga ggcctcttc tttgttcagt tcgagaaggg 480  
 cgagagctac tttcacctgc acgttctggt cgagaccacg ggggtcaagt ccatggtgct 540  
 50 aggccgcttc ctgagtcaga ttcgggaaaa gcttggtcca gaccatctac ccgcggggtc 600  
 gagccccacc ttgcccaact ggttcgcggt gaccaaagac gcggtaatgg cgccggcggg 660  
 ggggaacaag gtggtggacg agtgctacat ccccaactac ctctgcca agactcagcc 720  
 55 cgagctgcag tgggcgtgga ctaacatgga ggagtatata agcgcgtgct tgaacctggc 780

EP 1 310 571 A2

5	cgagcgcaaa cggctcgtgg cgcagcacct gaccacgctc agccagacgc aggagcagaa	840
	caaggagaaat ctgaacccca attctgacgc gcccgatgc aggtcaaaaa cctccgcgcg	900
	ctatatggag ctggtcgggt ggctgggtga ccggggcatc acctccgaga agcagtggat	960
10	ccaggaggac caggcctcgt acatctcctt caacgccgcc tccaactcgc ggtcccagat	1020
	caaggccgcg ctggacaatg ccggcaagat catggcgctg accaaatccg cggccgacta	1080
	cctgggtggg cctcgcctgc ccgcggacat taccagaaac cgcattctacc gcacccctgc	1140
15	tctcaacggc tacgacctg cctacgccg ctcgcgtctt ctcggctggg ctcagaaaaa	1200
	gttcgggaaa cgcaacacca tctggctggt tggaccgcc accaccggca agaccaacat	1260
	tgcggaagcc atcgccacg ccgtgccctt ctacggctgc gtcaactgga ccaatgagaa	1320
20	ctttcccttc aatgattgcg tcgacaagat ggtgatctgg tgggaggagg gcaagatgac	1380
	ggccaaggtc gtggagtccg ccaaggccat tctcggcggc agcaaggctc gcgtggacca	1440
	aaagtgcagg tcgtccgcc agatcgaccc caccgccgtg atcgtcacct ccaacaccaa	1500
25	catgtgcgcc gtgattgacg ggaacagcac caccctcgag caccagcagc ctctccagga	1560
	ccggatgttt aagttcgaac tcaccgccg tctggagcac gactttggca aggtgacaaa	1620
	gcaggaagtc aaagagttct tccgctgggc cagtgatcac gtgaccgagg tggcgcatga	1680
30	gttttacgtc agaaagggcg gagccagcaa aagaccgcc cccgatgacg cggataaaag	1740
	cgagcccaag cgggcctgcc cctcagtcgc ggatccatcg acgtcagacg cggaaggagc	1800
	tccggtggac tttgccgaca ggtaccaaaa caaatgttct cgtcacgcgg gcattgttca	1860
35	gatgctgttt ccctgcaaaa cgtgcgagag aatgaatcag aatttcaaca tttgcttcac	1920
	acacggggtc agagactgct cagagtgttt ccccgccgtg tcagaatctc aaccggctgt	1980
	cagaaagagg acgtatcgga aactctgtgc gattcatcat ctgctggggc gggctcccga	2040
40	gattgcttgc tcggcctgcg atctgggtcaa cgtggacctg gatgactgtg tttctgagca	2100
	ataaatgact taaaccagggt atggctgccg atggttatct tccagattgg ctcgaggaca	2160
	acctctctga gggcattcgc gagtgggtgg cgctgaaacc tggagccccg aagcccaaaag	2220
45	ccaaccagca aaagcaggac gacggccggg gtctggtgct tcctggctac aagtacctcg	2280
	gacccttcaa cggactcgac aagggggagc ccgtcaacgc ggcggacgca gcggccctcg	2340
	agcacgacaa ggcctacgac cagcagctgc aggcgggtga caatccgtac ctgcggtata	2400
50	accacgccga cggcgagttt caggagcgtc tgcaagaaga tacgtctttt gggggcaacc	2460
	tcgggcgagc agtcttccag gccaaagaag gggttctcga acctctcggg ctggttgagg	2520
	aaggcgctaa gacggctcct ggaagaaga gaccggtaga gccatcacc cagcgttctc	2580
55	cagactcctc tacgggcacg ggcaagaaag gccaacagcc cgccagaaaa agactcaatt	2640
	ttggtcagac tggcgactca gagtcagttc cagaccctca acctctcgga gaacctccag	2700

# EP 1 310 571 A2

5 cagcgccctc tgggtgtggga cctaatacaa tggctgcagg cgggtggcgca ccaatggcag 2760  
 acaataacga aggcgccgac ggagtgggta gttcctcggg aaattggcat tgcgattcca 2820  
 catggctggg cgacagagtc atcaccacca gcaccggaac ctgggccctg cccacctaca 2880  
 acaaccacct ctacaagcaa atctccaacg ggacatcggg aggagccacc aacgacaaca 2940  
 10 cctacttcgg ctacagcacc ccctgggggt attttgactt taacagattc cactgccact 3000  
 tttcaccacg tgactggcag cgactcatca acaacaactg gggattccgg cccaagagac 3060  
 tcagcttcaa gctcttcaac atccagggtca aggaggtcac gcagaatgaa ggcaccaaga 3120  
 15 ccatcgccaa taacctcacc agcaccatcc aggtgtttac ggactcggag taccagctgc 3180  
 cgtacgttct cggctctgcc caccagggct gcctgcctcc gttcccggcg gacgtgttca 3240  
 tgattcccca gtacggctac ctaacactca acaacggtag tcaggccgtg ggacgctcct 3300  
 20 ccttctactg cctggaatac ttctcttcgc agatgctgag aaccggcaac aacttccagt 3360  
 ttacttacac cttcgaggac gtgcctttcc acagcagcta cgcccacagc cagagcttgg 3420  
 accggctgat gaatcctctg attgaccagt acctgtacta cttgtctcgg actcaaacaa 3480  
 25 caggaggcac ggcaaatacg cagactctgg gcttcagcca aggtgggcct aatacaatgg 3540  
 ccaatcaggc aaagaactgg ctgccaggac cctgttaccg ccaacaacgc gtctcaacga 3600  
 caaccgggca aaacaacaat agcaactttg cctggactgc tgggaccaa taccatctga 3660  
 30 atggaagaaa ttcatgtggc aatcctggca tcgctatggc aacacacaaa gacgacgagg 3720  
 agcgtttttt tccagtaac gggatcctga tttttggcaa acaaaatgct gccagagaca 3780  
 atgcggatta cagcgatgtc atgctcacca gcgaggaaga aatcaaaacc actaacctg 3840  
 35 tggctacaga ggaatacggc atcgtggcag ataacttgca gcagcaaac acggctcctc 3900  
 aaattggaac tgtcaacagc cagggggcct taccgggtat ggtctggcag aaccgggacg 3960  
 tgtacctgca ggggtccatc tgggccaaga ttcttcacac ggacggcaac ttccacccgt 4020  
 40 ctccgctgat gggcggtttt ggcctgaaac atcctccgcc tcagatcctg atcaagaaca 4080  
 cgcctgtacc tgcggatcct ccgaccacct tcaaccagtc aaagctgaac tctttcatca 4140  
 cgcaatacag caccggacag gtcagcgtgg aaattgaatg ggagctgcag aaggaaaaca 4200  
 45 gcaagcgtg gaaccccgag atccagtaca cctccaacta ctacaaatct acaagtgtgg 4260  
 actttgctgt taatacagaa ggcgtgtact ctgaaccccg ccccatggc acccgttacc 4320  
 tcacccgtaa tctgtaattg cctgttaatc aataaaccgg ttgattcgtt tcagttgaac 4380  
 50 tttggtctct gcg 4393

<210> 5  
 <211> 4385  
 <212> DNA  
 <213> adeno-associated virus serotype 9

55

EP 1 310 571 A2

<400> 5  
5 cagagagggga gtggccaact ccatactag gggtaatcgc gaagcgcctc ccacgctgcc 60  
gcgtcagcgc tgacgtagat tacgtcatag gggagtggtc ctgtattagc tgtcacgtga 120  
gtgcttttgc gacattttgc gacaccacat ggccatttga ggtatatatg gccgagtgag 180  
10 cgagcaggat ctccattttg accgcgaaat ttgaacgagc agcagccatg ccgggcttct 240  
acgagattgt gatcaagggtg ccgagcgacc tggacgagca cctgccgggc atttctgact 300  
cttttgtgaa ctgggtggcc gagaaggaaat gggagctgcc cccgatttct gacatcgatc 360  
ggaatctgat cgagcaggca cccctgaccg tggccgagaa gctgcagcgc gacttctggtg 420  
15 tccaatggcg ccgctgagtg aaggcccccg agggcctctt ctttgttcag ttcgagaagg 480  
gcgagagcta ctttcacctg cacgttcttg tgcagaccac gggggtcaag tccatggtgc 540  
taggcccgtt cctgagtcag attcgggaga agctggtcca gaccatctac cgcgggatcg 600  
20 agccgaccct gcccaactgg ttgcgggtga ccaagacgcg taatggcgcc ggcgggggga 660  
acaaggtggt ggacgagtg ccatcccca actacctct gcccaagact cagcccagc 720  
tgcagtgggc gtggactaac atggaggagt atataagcgc gtgcttgaac ctggccgagc 780  
25 gcaaacggct cgtggcgag cacctgacct acgtcagcca gacgcaggag cagaacaagg 840  
agaatctgaa cccaattct gacgcgccc tgatcaggtc aaaaacctc gcgcgtaca 900  
tggagctggt cgggtggctg gtggaccggg gcatcacctc cgagaagcag tggatccagg 960  
30 aggaccaggc ctctacatc tcttcaacg ccgcctcaa ctgcgggtcc cagatcaagg 1020  
ccgcgtgga caatgccggc aagatcatgg cgctgaccaa atccgcgcc gactacctg 1080  
taggcccctt acttccggtg gacattacgc agaaccgat ctaccgatc ctgcagctca 1140  
35 acggctacga cctgectac gccggtccg tctttctcgg ctgggcacaa aagaagtctg 1200  
ggaaacgcaa caccatctg ctgtttgggc cggccaccac gggaaagacc aacatcgag 1260  
aagccattgc ccacgcgtg ccttctacg gctgcgtcaa ctggaccaat gagaactttc 1320  
40 ccttcaacga ttgcgtcgac aagatggtga tctggtggga ggagggcaag atgacggcca 1380  
aggtcgtgga gtccgccaa gcaattctcg gcggcagcaa ggtgcgctg gacaaaaagt 1440  
gcaagtcgtc ccccagatc gacccactc ccgtgatcgt cacctccaac accaactgt 1500  
45 gcgcctgat tgacgggaac agcaccacct tcgagacca gcagcctctc caggaccgga 1560  
tgtttaagtt cgaactcacc cgccgtctgg agcacgactt tggcaagggtg acaaagcagg 1620  
aagtcaaaga gttcttcgc tgggccagt atcacgtgac cgaggtggcg catgagtttt 1680  
50 acgtcagaaa gggcggagcc agcaaaagac ccgccccga tgacgcggat aaaagcgagc 1740  
ccaagcgggc ctgccctca gtgcggatc catcgacgtc agacgcggaa ggagctccg 1800  
tggactttgc cgacaggtac caaaacaaat gttctcgtca cgcgggcatg cttcagatgc 1860  
55 tgcttcctg caaacgtgc gagagaatga atcagaattt caacatttgc ttcacacag 1920

EP 1 310 571 A2

5	gggtcagaga ctgctcagag tgtttccccg gcgtgtcaga atctcaaccg gtcgtcagaa	1980
	agaggacgta tcggaaactc tgtgcgattc atcatctgct ggggcgggct cccgagattg	2040
	cttgctcggc ctgcgatctg gtcaacgtgg acctggatga ctgtgtttct gagcaataaa	2100
10	tgacttaaac caggtatggc tgccgatggc tatcttccag attggctcga ggacaacctc	2160
	tctgagggca ttgcgagtg gtgggcgctg aaacctggag ccccgaaagc caaagccaac	2220
	cagcaaaagc aggacgacg ccgggggtctg gtgcttctctg gctacaagta cctcggaccc	2280
15	ttcaacggac tcgacaaggg ggagcccgtc aacgcggcgg acgcagcggc cctcagacac	2340
	ggcaaggcct acgaccagca gctgcaggcg ggtgacaatc cgtacctgag gtataaccac	2400
	gccgacgccg agtttcagga gcgtctgcaa gaagatacgt cttttggggg caacctcggg	2460
20	cgagcagtct tccaggccaa gaagcggggt ctcgaacctc tcgggtctggg tgaggaaggc	2520
	gctaagacgg ctccctggaaa gaagagaccg gtagagccat caccacagcg ttctccagac	2580
	tcctctacgg gcatcgccaa gaaaggccaa cagcccgcca gaaaaagact caattttggt	2640
25	cagactggcg actcagagtc agttccagac cctcaacctc tcggagaacc tccagcagcg	2700
	ccctctgggt tgggacctaa tacaatggct gcaggcgggtg gcgcaccaat ggcagacaat	2760
	aacgaaggcg ccgacggagt gggtaattcc tcgggaaatt ggcattgcga ttccacatgg	2820
30	ctgggggaca gagtcatcac caccagcacc cgaacctggg cattgcccac ctacaacaac	2880
	cacctctaca agcaaactct caatggaaca tcgggaggaa gcaccaacga caaacctac	2940
	tttggctaca gcacccccctg ggggtatttt gacttcaaca gattccactg ccacttctca	3000
35	ccacgtgact ggcagcgact catcaacaac aactggggat tccggccaaa gagactcaac	3060
	ttcaagctgt tcaacatcca ggtcaaggag gttacgacga acgaaggcac caagaccatc	3120
	gccataaacc ttaccagcac cgtccaggtc tttacggact cggagtacca gctaccgtac	3180
40	gtcctaggct ctgcccacca aggatgcctg ccaccgtttc ctgcagacgt cttcatgggt	3240
	cctcagtagc gctacctgac gctcaacaat ggaagtcaag cgtaggagc ttcttctttc	3300
	tactgtctgg aatacttccc ttctcagatg ctgagaaccg gcaacaactt tcagttcagc	3360
45	tacactttcg aggacgtgcc tttccacagc agctacgcac acagccagag tctagatcga	3420
	ctgatgaacc ccctcatcga ccagtaccta tactacctgg tcagaacaca gacaactgga	3480
	actgggggaa ctcaaacttt ggcattcagc caagcaggcc ctagctcaat ggccaatcag	3540
50	gctagaaact gggtagccgg gccttgctac cgtcagcagc gcgtctccac aaccaccaac	3600
	caaaataaca acagcaactt tgcgtggagc ggagctgcta aattcaagct gaacgggaga	3660
	gactcgctaa tgaatcctgg cgtggctatg gcatcgaca aagacgacga ggaccgcttc	3720
	tttccatcaa gtggcgctct catatttggc aagcaaggag ccgggaacga tggagtcgac	3780
55	tacagccagg tgctgattac agatgaggaa gaaattaaag ccaccaaccc tgtagccaca	3840

# EP 1 310 571 A2

5 gaggaatacgc gagcagtggc catcaacaac caggccgcta acacgcaggc gcaaactgga 3900  
 cttgtgcata accagggagt tattcctggg atggtctggc agaaccggga cgtgtacctg 3960  
 cagggcccta tttgggctaa aatacctcac acagatggca actttcaccc gtctcctctg 4020  
 atgggtggat ttggactgaa acaccacact ccacagattc taattaaaaa tacaccagtg 4080  
 10 ccggcagatc ctctcttac cttcaatcaa gccaaagtga actctttcat cagcagtagc 4140  
 agcacgggac aagtcagcgt ggaaatcgag tgggagctgc agaaagaaaa cagcaagcgc 4200  
 15 tggaaatccag agatccagta tacttcaaac tactacaaat ctacaaatgt ggactttgct 4260  
 gtcaataacca aaggtgttta ctctgagcct cgccccattg gtactcgtta cctcaccctg 4320  
 aatttgtaat tgctgtttaa tcaataaacc ggttaattcg tttcagttga actttggctc 4380  
 20 ctgcg 4385

<210> 6  
 <211> 4718  
 <212> DNA  
 <213> adeno-associated virus serotype 1

25 <400> 6  
 ttgccactc cctctctgcg cgctcgctcg ctcggtgggg cctgcggacc aaaggtccgc 60  
 agacggcaga gctctgctct gccggcccca ccgagcgagc gagcgcgag agagggagtg 120  
 30 ggcaactcca tctactaggg taatcgcgaa gcgcctccca cgctgccgag tcagcgctga 180  
 cgtaaattac gtcataaggg agtggctctg tattagctgt cagctgagtg cttttgcgac 240  
 attttgcgac accacgtggc catttagggg atatatggcc gactgagcga gcaggatctc 300  
 35 cattttgacc gcgaaatttg aacgagcagc agccatgccg ggcttctacg agatcgtgat 360  
 caaggtgccg agcgacctgg acgagcacct gccgggcatt tctgactcgt ttgtgagctg 420  
 ggtggccgag aaggaatggg agctgcccc ggattctgac atgcatctga atctgattga 480  
 40 gcaggcacc ctagacctgg ccgagaagct gcagcgcgac ttcttggtcc aatggcgccg 540  
 cgtgagtaag gccccggagg ccctcttctt tgttcagttc gagaagggcg agtcctactt 600  
 ccacctccat attctggtgg agaccacggg ggtcaaattc atggtgctgg gccgcttcct 660  
 45 gactcagatt agggacaagc tgggtgcagac catctaccgc gggatcgagc cgacctgccc 720  
 caactggttc gcggtgacca agacgcgtaa tggcgccgga ggggggaaca aggtggtgga 780  
 cgagtgtctac atccccaaact acctcctgcc caagactcag cccgagctgc agtgggctg 840  
 50 gactaacatg gaggagtata taagcgcttg tttgaacctg gccgagcgca aacggctcgt 900  
 ggcgagcac ctagccagac ccaggagcag aacaaggaga atctgaaccc 960  
 caattctgac gcgcctgtca tccggtcaaa aacctccgag cgctacatgg agctggctcg 1020  
 55 gtggctggtg gaccggggca tcacctccga gaagcagtg atccaggagg accaggcctc 1080

EP 1 310 571 A2

	gtacatctcc ttcaacgccg cttccaactc gcggtcccag atcaaggccg ctctggacaa	1140
5	tgccggcaag atcatggcgc tgaccaaato cgcgcccgac tacctggtag gccccgctcc	1200
	gccccgggac attaaaacca accgcatcta ccgcatcctg gagctgaacg gctacgaacc	1260
	tgcctacgcc ggctccgtct ttctcggctg ggcccagaaa aggttcggga agcgcaacac	1320
10	catctggctg tttggggccg ccaccacggg caagaccaac atcgcggaag ccatcgccca	1380
	cggcgtgccc ttctacggct gcgtcaactg gaccaatgag aactttccct tcaatgattg	1440
	cgtcgacaag atggtgatct ggtgggagga gggcaagatg acggccaagg tcgtggagtc	1500
15	cgccaaggcc attctcggcg gcagcaaggc gcgcgtggac caaaagtgc agtcgtccgc	1560
	ccagatcgac cccacccccg tgatcgtcac ctccaacacc aacatgtgcg ccgtgattga	1620
	cgggaacagc accaccttcg agcaccagca gccgttgacg gaccggatgt tcaaatttga	1680
20	actcaccgc cgtctggagc atgactttgg caaggtgaca aagcaggaag tcaaagagtt	1740
	cttccgctgg gcgcaggatc acgtgaccga ggtggcgcat gagttctacg tcagaaaggg	1800
	tggagccaac aaaagacccg ccccgatga cgcgataaa agcgagccca agcgggctg	1860
25	cccctcagtc gcggatccat cgacgtcaga cgcggaagga gctccggtgg actttgccga	1920
	caggtaccaa aacaaatgtt ctcgtcacgc gggcatgctt cagatgctgt ttccctgcaa	1980
	gacatgcgag agaatgaatc agaatttcaa catttgcttc acgcacggga cgagagactg	2040
30	ttcagagtgc ttccccggcg tgtcagaatc tcaaccggtc gtcagaaaga ggacgtatcg	2100
	gaaactctgt gccattcatc atctgctggg gcgggctccc gagattgctt gctcggcctg	2160
	cgatctggtc aacgtggacc tggatgactg tgtttctgag caataaatga cttaaaccag	2220
35	gtatggctgc cgatggttat cttccagatt ggctcgagga caacctctct gagggcattc	2280
	gcgagtggcg ggacttgaaa cctggagccc cgaagcccaa agccaaccag caaaagcagg	2340
	acgacggccg gggctctggcg cttcctggct acaagtacct cggacccttc aacggactcg	2400
40	acaaggggga gcccgtaac gcggcgagcg cagcggccct cgagcacgac aaggcctacg	2460
	accagcagct caaagcgggt gacaatccgt acctgcggta taaccacgcc gacgccgagt	2520
	ttcaggagcg tctgcaagaa gatacgtctt ttgggggcaa cctcgggcga gcagtcttcc	2580
45	aggccaagaa gcgggttctc gaacctctcg gtctggttga ggaaggcgtc aagacggctc	2640
	ctggaaagaa acgtccggtg gagcagtcgc cacaagagcc agactcctcc tcgggcatcg	2700
	gcaagacagg ccagcagccc gctaaaaaga gactcaattt tggtcagact ggcgactcag	2760
50	agtcagtccc cgatccacaa cctctcggag aacctccagc aacccccgct gctgtgggac	2820
	ctactacaat ggcttcaggc ggtggcgac caatggcaga caataacgaa ggcgccgacg	2880
	gagtgggtaa tgcctcagga aattggcatt gcgattccac atggctgggc gacagagtca	2940
55	tcaccaccag caccgcacc tgggccttgc ccacctacaa taaccacctc tacaagcaaa	3000

# EP 1 310 571 A2

	tctccagtgc ttcaacgggg gccagcaacg acaaccacta cttcgggtac agcaccacct	3060
5	gggggtatatt tgatttcaac agattccact gccacttttc accacgtgac tggcagcgac	3120
	tcatcaacaa caattgggga ttccggccca agagactcaa cttcaaactc ttcaacatcc	3180
	aagtcaagga ggtcacgacg aatgatggcg tcacaaccat cgctaataac cttaccagca	3240
10	cgggttcaagt cttctcggac tcggagtacc agcttcgta cgtcctcggc tctgcgcacc	3300
	agggtgcct ccctccgttc ccggcggacg tgttcatgat tccgcaatac ggctacctga	3360
	cgctcaacaa tggcagccaa gccgtgggac gttcctcctt ttactgcctg gaatatttcc	3420
15	cttctcagat gctgagaacg ggcaacaact ttaccttcag ctacaccttt gaggaagtgc	3480
	ctttccacag cagctacgcy cacagccaga gcctggaccg gctgatgaat cctctcatcg	3540
	accaatacct gtattacctg aacagaactc aaaatcagtc cggaaagtgc caaaacaagg	3600
20	acttgctgtt tagccgtggg tctccagctg gcatgtctgt tcagcccaaa aactggctac	3660
	ctggaccctg ttatcggcag cagcgcgttt ctaaaacaaa aacagacaac aacaacagca	3720
	attttacctg gactggtgct tcaaaatata acctcaatgg gcgtgaatcc atcatcaacc	3780
25	ctggcactgc tatggcctca cacaaagacg acgaagacaa gttctttccc atgagcggtg	3840
	tcatgatttt tggaaaagag agcgcgggag cttcaaacac tgcattggac aatgtcatga	3900
	ttacagacga agaggaaatt aaagccacta acctgtggc caccgaaaga tttgggaccg	3960
30	tggcagtcaa tttccagagc agcagcacag acctgcgac cggagatgtg catgctatgg	4020
	gagcattacc tggcatggtg tggcaagata gagacgtgta cctgcagggc cccatttggg	4080
	ccaaaattcc tcacacagat ggacactttc acccgtctcc tcttatgggc ggctttggac	4140
35	tcaagaacct gcctcctcag atcctcatca aaaacacgcc tgttcctgcy aatcctccgg	4200
	cggagttttc agctacaaag tttgcttcat tcatcaccca atactccaca ggacaagtga	4260
	gtgtggaaat tgaatgggag ctgcagaaag aaaacagcaa gcgctggaat cccgaagtgc	4320
40	agtacacatc caattatgca aaatctgcc aagttgattt tactgtggac aacaatggac	4380
	tttatactga gcctcgcccc attggcaccg gttaccttac ccgtccctg taattacgtg	4440
	ttaatcaata aaccggttga ttcgtttcag ttgaactttg gtctcctgtc cttcttatct	4500
45	tatcggttac catggttata gcttacacat taactgcttg gttgcgcttc gcgataaaag	4560
	acttacgtca tcgggttacc cctagtgatg gaggttgccca ctccctctct gcgcgctcgc	4620
	tcgctcggtg gggcctgcgg accaaaggtc cgcagacggc agagctctgc tctgccggcc	4680
50	ccaccgagcg agcgagcgcg cagagaggga gtgggcaa	4718

<210> 7  
 <211> 4675  
 <212> DNA  
 <213> adeno-associated virus serotype 2

# EP 1 310 571 A2

<400> 7  
5 ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggctcgcc 60  
cgacgcccg gctttgcccg ggcgccctca gtgagcgagc gagcgcgag agagggagtg 120  
gccaaactcca tactagggg ttcctggagg ggtggagtcg tgacgtgaat tacgtcatag 180  
10 ggttagggag gtctgtatt agaggtcagc tgagtgtttt gcgacatttt gcgacaccat 240  
gtggtcacgc tgggtattta agcccgagtg agcacgcagg gtctccattt tgaagcggga 300  
ggtttgaacg cgacgccgc atgcccgggt tttacgagat tgtgattaag gtccccagcg 360  
15 accttgacgg gcatctgcc ggcatctctg acagctttgt gaactgggtg gccgagaagg 420  
aatgggagtt gccgccagat tctgacatgg atctgaatct gattgagcag gcacccctga 480  
ccgtggccga gaagctgcag cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc 540  
cggaggccct tttctttgtg caatttgaga agggagagag ctacttccac atgcacgtgc 600  
20 tcgtggaaac caccggggtg aaatccatgg ttttgggacg tttcctgagt cagattcgcg 660  
aaaaactgat tcagagaatt taccgcgga tcgagccgac tttgccaaac tggttcgcg 720  
tcacaaagac cagaaatggc gccggaggcg ggaacaagggt ggtggatgag tgctacatcc 780  
25 ccaattactt gctcccaaa acccagcctg agctccagtg gccgtggact aatatggaac 840  
agtatttaag cgctgtttg aatctcacgg agcgtaaacg gttggtggcg cagcatctga 900  
cgcacgtgtc gcagacgcag gagcagaaca aagagaatca gaatcccaat tctgatgcgc 960  
30 cggatgatcag atcaaaaact tcagccaggt acatggagct ggtcgggtgg ctctgggaca 1020  
aggggattac ctcggaag cagtggatcc aggaggacca ggctcctac atctccttca 1080  
atgcccctc caactcgcg tcccaaatca aggtgcctt ggacaatgcg ggaagatta 1140  
35 tgagcctgac taaaaccgcc ccgactacc tgggtggcca gcagcccggt gaggacattt 1200  
ccagcaatcg gatttataaa attttgaac taaacgggta cgatcccaa tatgcggtt 1260  
ccgtctttct gggatgggcc acgaaaaagt tcggcaagag gaacaccatc tggctgtttg 1320  
40 ggcctgcaac taccgggaag accaaccatc cggaggccat agcccacact gtgcccttct 1380  
acgggtgcgt aaactggacc aatgagaact ttcccttcaa cgactgtgtc gacaagatgg 1440  
tgatctggtg ggaggagggg aagatgaccg ccaaggctgt ggagtcggcc aaagccattc 1500  
45 tcggaggaag caaggtgcgc gtggaccaga aatgcaagtc ctcgcccag atagaccga 1560  
ctcccgatg cgtcacctcc aacaccaaca tgtgcgccgt gattgacggg aactcaacga 1620  
ccttcgaaca ccagcagccg ttgcaagacc ggatgttcaa atttgaactc acccgccgtc 1680  
50 tggatcatga ctttgggaag gtcaccaagc aggaagtcaa agactttttc cgggtgggcaa 1740  
aggatcacgt ggttgaggtg gagcatgaat tctactgcaa aaaggtgga gccaaagaaa 1800  
gacccgcccc cagtgcgca gatataagt agcccaaacg ggtgcgcgag tcagttgcgc 1860  
55 agccatcgac gtcagacgc gaagcttca tcaactacgc agacaggtag caaaacaaat 1920

# EP 1 310 571 A2

5	gttctcgtca cgtgggcatg aatctgatgc tgtttccctg cagacaatgc gagagaatga	1980
	atcagaattc aaatatctgc ttcactcacg gacagaaaga ctgttttagag tgctttcccg	2040
	tgtcagaatc tcaaccogtt tctgtcgtca aaaaggcgta tcagaaactg tgctacattc	2100
10	atcatatcat gggaaggtg ccagacgctt gcactgcctg cgatctggtc aatgtggatt	2160
	tggatgactg catctttgaa caataaatga tttaaatcag gtatggctgc cgatgggttat	2220
	cttccagatt ggctcgagga cactctctct gaaggaataa gacagtgggtg gaagctcaaa	2280
15	cctggcccac caccaccaa gcccgcagag cggcataagg acgacagcag gggctctgtg	2340
	cttcctgggt acaagtacct cggacccttc aacggactcg acaagggaga gccgggtcaac	2400
	gaggcagacg ccgcggccct cgagcacgta caaagcctac gaccggcagc tcgacagcgg	2460
20	agacaacccg tacctcaagt acaaccacgc cgacgcggag tttcaggagc gccttaaaga	2520
	agatacgtct tttgggggca acctcggacg agcagtcttc caggcgaaaa agaggggttct	2580
	tgaacctctg ggcttggttg aggaacctgt taagacggct ccgggaaaaa agaggccggt	2640
25	agagcactct cctgtggagc cagactcctc ctcggaacc ggaagggcg gccagcagcc	2700
	tgcaagaaaa agattgaatt ttggtcagac tggagacgca gactcagtag ctgaccccca	2760
	gcctctcgga cagccaccag cagccccctc tgggtctggga actaatacga tggctacagg	2820
30	cagtggcgca ccaatggcag acaataacga gggcgccgac ggagtgggta attcctccgg	2880
	aaattggcat tgcgattcca catggatggg cgacagagtc atcaccacca gcacccgaac	2940
	ctgggcccctg cccacctaca acaaccacct ctacaaacaa atttccagcc aatcaggagc	3000
35	ctcgaacgac aatcactact ttggtacag cacccttgg ggggtattttg acttcaacag	3060
	attccactgc cacttttcac cactgtactg gcaaagactc atcaacaaca actggggatt	3120
	ccgacccaag agactcaact tcaagctctt taacattcaa gtcaaagagg tcacgcagaa	3180
40	tgacggtacg acgacgattg ccaataacct taccagcagc gttcagggtgt ttactgactc	3240
	ggagtaccag ctcccgtacg tctcggctc ggcgcacaa ggatgcctcc cgcggttccc	3300
	agcagacgtc ttcattggtg cacagtatgg atacctcacc ctgaacaacg ggagtcaggc	3360
45	agtaggacgc tcttcatttt actgcctgga gtactttcct tctcagatgc tgcgtaccgg	3420
	aaacaacttt accttcagct acacttttga ggacgttcct ttccacagca gctacgctca	3480
	cagccagagt ctggaccgtc tcatgaatcc tctcatcgac cagtacctgt attacttgag	3540
50	cagaacaaac actccaagtg gaaccaccac gcagtcaagg cttcagtttt ctcaggcccg	3600
	agcgagtgac attcgggacc agtctaggaa ctggcttcct ggaccctgtt accgccagca	3660
	gcgagtatca aagacatctg cggataacaa caacagtga tactcgtgga ctggagctac	3720
55	caagtaccac ctcaatggca gagactctct ggtgaatccg gccatggcaa gccacaagga	3780
	cgatgaagaa aagttttttc ctcagagcgg ggttctcatc tttgggaagc aaggctcaga	3840

# EP 1 310 571 A2

5 gaaaacaaat gtgaacattg aaaaggtcat gattacagac gaagaggaaa tcggaacaac 3900  
 caatcccgtg gctacggagc agtatggttc tgtatctacc aacctccaga gaggcaacag 3960  
 acaagcagct accgcagatg tcaacacaca aggcgttctt ccaggcatgg tctggcagga 4020  
 10 cagagatgtg taccttcagg ggcccattcg ggcaaagatt ccacacacgg acggacattt 4080  
 tcaccctctt cccctcatgg gtggattcgg acttaaacac cctcctccac agatttctcat 4140  
 caagaacacc ccggtacctg cgaatccttc gaccaccttc agtgcggaag agtttgcttc 4200  
 cttcatcaca cagtactcca cgggacacgg tcagcgtgga gatcgagtgg gagctgcaga 4260  
 15 agggaaacag caaacgctgg aatcccgaaa ttcagtacac ttccaactac aacaagtctg 4320  
 ttaatcgtgg acttaccgtg gataactaat gcggtgtattc agagcctcgc cccattggca 4380  
 ccagatacct gactcgtaat ctgtaattgc ttgttaatca ataaaccgtt taattcgttt 4440  
 20 cagttgaact ttggtctctg cgtatttctt tcttatctag tttccatggc tacgtagata 4500  
 agtagcatgg cgggttaatc attaaactaca aggaacccct agtgatggag ttggccactc 4560  
 cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc cgacgcccgg 4620  
 25 gctttgcccg ggcgccctca gtgagcgagc gagcgcgag agaggagtg gccaa 4675

<210> 8  
 <211> 4726  
 <212> DNA  
 <213> adeno-associated virus serotype 3

30 <400> 8  
 ttggccactc cctctatgcg cactcgctcg ctcggtgggg cctggcgacc aaaggtcgcc 60  
 35 agacggacgt gctttgcacg tccggcccca ccgagcgagc gagtgcgcat agaggagtg 120  
 gccaaactcca tctactagagg tatggcagtg acgtaacgag aagcgcgcga agcgagacca 180  
 cgcctaccag ctgcgtcagc agtcaggtga ccccttttgcg acagtttgcg acaccacgtg 240  
 40 gccgctgagg gtatatattc tcgagtgagc gaaccaggag ctccattttg accgcgaaat 300  
 ttgaacgagc agcagccatg ccgggggttct acgagattgt cctgaaggtc ccgagtgacc 360  
 tggacgagcg cctgccgggc atttctaact cgtttgttaa ctgggtggcc gagaaggaat 420  
 45 gggacgtgcc gccggttctt gacatggatc cgaatctgat tgagcaggca cccctgaccg 480  
 tggccgaaaa gcttcagcgc gagttccttg tggagtggcg ccgctgagt aaggccccgg 540  
 aggccctctt ttttgtccag ttcgaaaagg gggagacctt cttccacctg cactgtctga 600  
 50 ttgagaccat cggggtcaaa tccatgggtg tcggccgcta cgtgagccag attaaagaga 660  
 agctgggtgac ccgcatctac cgcggggtcg agccgcagct tccgaactgg ttcgcggtga 720  
 ccaaaacgag aaatggcgcc gggggcgga acaaggtggt ggacgactgc tacatcccca 780  
 55 actacctgct ccccaagacc cagcccagc tccagtgggc gtggactaac atggaccagt 840

EP 1 310 571 A2

	atttaagcgc ctgtttgaat ctgcgcgagc gtaaaccgct ggtggcgcag catctgacgc	900
5	acgtgtcgca gacgcaggag cagaacaaag agaatcagaa ccccaattct gacgcgccgg	960
	tcattcaggtc aaaaacctca gccaggtaga tggagctggt cgggtggctg gtggaccgcg	1020
	ggatcacgtc agaaaagcaa tggattcagg aggaccaggc ctgttacatc tccttcaacg	1080
10	ccgcctccaa ctgcgcgtcc cagatcaagg ccgcgctgga caatgcctcc aagatcatga	1140
	gcctgacaaa gacggctccg gactacctgg tgggcagcaa cccgccggag gacattacca	1200
	aaaatcggat ctaccaaadc ctggagctga acgggtacga tccgcagtac gcggcctccg	1260
15	tcttcctggg ctgggcgcaa aagaagtctg ggaagaggaa caccatctgg ctctttgggc	1320
	cggccacgac gggtaaaacc aacatcgcgg aagccatcgc ccacgccgtg cccttctacg	1380
	gctgcgtaaa ctggaccaat gagaactttc ccttcaacga ttgcgtcgac aagatggtga	1440
20	tctggtggga ggagggcaa atgacggcca aggtcgtgga gagcgccaag gccattctgg	1500
	gcggaagcaa ggtgcgcgtg gacaaaaagt gcaagtcacg ggcccagatc gaaccactc	1560
	ccgtgatcgt cactccaac accaacaatg gcgccgtgat tgacgggaac agcaccacct	1620
25	tcgagcatca gcagccgtg caggaccgga tgtttgaatt tgaacttacc cgccgtttgg	1680
	accatgactt tgggaaggtc accaaacagg aagtaaagga ctttttccgg tgggcttccg	1740
	atcacgtgac tgacgtggct catgagttct acgtcagaaa ggggtggagct aagaaacgcc	1800
30	ccgcctccaa tgacgcggat gtaagcgagc caaacggga gtgcacgtca cttgcgcagc	1860
	cgacaacgtc agacgcggaa gcaccggcgg actacgcgga caggtagcaa aacaaatggt	1920
	ctcgtcacgt gggcatgaat ctgatgcttt ttccctgtaa aacatgagag agaataatc	1980
35	aaatttccaa tgtctgtttt acgcatggtc aaagagactg tggggaatgc ttccctggaa	2040
	tgtcagaatc tcaaccggtt tctgtcgtca aaaagaagac ttatcagaaa ctgtgtccaa	2100
	ttcatcatat cctgggaagg gcacccgaga ttgcctgttc ggcctgcgat ttggccaatg	2160
40	tggacttgga tgactgtgtt tctgagcaat aaatgactta aaccaggatg ggctgctgac	2220
	ggttatcttc cagattggct cgaggacaac ctttctgaag gcattcgtga gtggtgggct	2280
	ctgaaacctg gagtccctca acccaaagcg aaccaacaac accaggacaa ccgtcggggt	2340
45	cttgtgtctc cgggttacaa atacctcgga cccggtaacg gactcgacaa aggagagccg	2400
	gtcaacgagg cggacgcggc agccctcgaa cagacaaaag cttacgacca gcagctcaag	2460
	gccggtgaca acccgtagct caagtacaac cagccgagc cagagtttca ggagcgtctt	2520
50	caagaagata cgtcttttgg gggcaacctt ggcagagcag tcttccaggc caaaaagagg	2580
	atccttgagc ctcttggtct ggttgaggaa gcagctaaaa cggctcctgg aaagaagggg	2640
	gctgtagatc agtctcctca ggaaccggac tcattcatctg gtgttgccaa atcgggcaaa	2700
55	cagcctgccg gaaaaagact aaatttcggt cagactggag actcagagtc agtcccagac	2760

EP 1 310 571 A2

	cctcaacctc tcggagaacc accagcagcc cccacaagtt tgggatctaa tacaatggct	2820
5	tcaggcggtg gcgcaccaat ggcagacaat aacgaggggtg ccgatggagt gggtaattcc	2880
	tcaggaaatt ggcattgcga ttcccaatgg ctgggcgaca gagtcatcac caccagcacc	2940
	agaacctggg ccctgcccac ttacaacaac catctctaca agcaaattct cagccaatca	3000
10	ggagcttcaa acgacaacca ctactttggc tacagcacc cttgggggta ttttgacttt	3060
	aacagattcc actgccactt ctcaccacgt gactggcagc gactcattaa caacaactgg	3120
	ggattccggc ccaagaaact cagcttcaag ctcttcaaca tccaagttag aggggtcacg	3180
15	cagaacgatg gcacgacgac tattgccaat aaccttacca gcacggttca agtgtttacg	3240
	gactcggagt atcagctccc gtacgtgctc gggtcggcgc accaaggctg tctcccgcg	3300
	tttccagcgg acgtcttcat ggtccctcag tatggatacc tcacctgaa caacggaagt	3360
20	caagcggtag gacgctcatc cttttactgc ctggagtact tcccttcgca gatgctaagg	3420
	actggaaata acttccaatt cagctatacc ttcgaggatg taccttttca cagcagctac	3480
	gctcacagcc agagtgttga tcgcttgatg aatcctctta ttgatcagta tctgtactac	3540
25	ctgaacagaa cgcaaggaa aacctctgga acaaccaacc aatcacggct gctttttagc	3600
	caggctgggc ctcagcttat gtctttgcag gccagaaatt ggctacctgg gccctgctac	3660
	cggcaacaga gactttcaaa gactgctaac gacaacaaca acagtaactt tccttgga	3720
30	gcggccagca aatatcatct caatggccgc gactcgtctg tgaatccagg accagctatg	3780
	gccagtcaca aggacgatga agaaaaattt ttccctatgc acggcaatct aatatttggc	3840
	aaagaaggga caacggcaag taacgcagaa ttagataatg taatgattac ggatgaagaa	3900
35	gagattcgtg ccaccaatcc tgtggcaaca gagcagtatg gaactgtggc aaataacttg	3960
	cagagctcaa atacagctcc cagcactgga actgtcaatc atcagggggc cttacctggc	4020
	atgggtgtgg aagatcgtga cgtgtacctt caaggacctt tctgggcaaa gattcctcac	4080
40	acggatggac actttcatcc ttctcctctg atgggaggct ttggactgaa acatccgcct	4140
	cctcaaatca tgatcaaaaa tactccggtg ccggcaaatc ctccgacgac tttcagcccg	4200
	gccaaagttg cttcatttat cactcagtac tccactggac aggtcagcgt ggaaattgag	4260
45	tgggagctac agaaagaaaa cagcaaactg tggaatccag agattcagta cacttccaac	4320
	tacaacaagt ctgttaatgt ggactttact gtagacacta atgggtgtta tagtgaacct	4380
	cgcctatttg gaaccgggta tctcacagca aacttgtgaa tcctgggtta tcaataaacc	4440
50	gtttaattcg tttcagttga actttggctc ttgtgcactt ctttatcttt atcttgtttc	4500
	catggtact gcgtagataa gcagcggcct gcggcgcttg cgcttcgagg tttacaactg	4560
	ctggttaata tttaactctc gccataacct tagtgatgga gttggccact ccctctatgc	4620
55	gcactcgtc gctcgggtgg gcctggcgac caaaggctgc cagacggagc tgctttgcac	4680

# EP 1 310 571 A2

gtccggcccc accgagcgag cgagtgcgca tagagggagt ggccaa 4726

5  
 <210> 9  
 <211> 3098  
 <212> DNA  
 <213> new AAV serotype, clone 42.2

10  
 <400> 9  
 gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60  
 gcgtcgacaa gatggtgatc tgggtgggagg agggcaagat gacggccaag gtcgtggagt 120  
 ccgccaaggc cattctcggc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcttccg 180  
 15  
 ccagatcga tcccaccccc gtgatcgta cttccaacac caacatgtgc gctgtgattg 240  
 acgggaacag caccaccttc gagcaccagc agccgttaca agaccgatg ttcaaatttg 300  
 aactcaccgg ccgtctggag cagcactttg gcaaggtgac aaagcaggaa gtcaaagagt 360  
 20  
 tcttccgctg ggcgaggat cactgtaccg aggtggcgca tgagttctac gtcagaaagg 420  
 gtggagccaa caagagaccc gcccccgatg acgcggataa aagcgagccc aagcgggcct 480  
 gccctcagt cgcgatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540  
 25  
 acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca 600  
 agacatgcga gagaatgaat cagaatttca acatttgctt cagcacggg accagagact 660  
 gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc 720  
 30  
 ggaaactctg tgccattcat catctgtctg ggcgggctcc cgagattgct tgctcggcct 780  
 gcgatctggt caacgtggac ctggatgacc gtgtttctga gcaataaatg acttaaacca 840  
 ggtatggctg ccgatggta tcttcagat tggctcgagg acaacctctc tgagggcatt 900  
 35  
 cgcgagtgtt gggacttgaa acctggagcc ccgaaacca aagccaacca gcaaaagcag 960  
 gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc 1020  
 gacaagggag agccggtcaa cgaggcagac gccgcggccc tcgagcacga caaggcctac 1080  
 40  
 gacaagcagc tcgagcaggg ggacaacccg tacctcaagt acaaccacgc cgacgccgag 1140  
 tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggag agcagtcttc 1200  
 caggccaaga agcgggttct cgaacctctc ggtctggtt aggaaggcgc taagacggct 1260  
 45  
 cctggaaaga agagacccat agaatcccc gactcctcca cgggcatcgg caagaaaggc 1320  
 cagcagcccc ctaaaaagaa gctcaacttt gggcagactg gcgactcaga gtcagtcccc 1380  
 gacccccaac ctctcggaga acctcccgcc gcgccctcag gtctgggatc tgggtacaatg 1440  
 50  
 gctgcaggcg gtggcgacc aatggcagac aataacgaag gcgccgacgg agtgggtaat 1500  
 gcctccgga attggcattg cgattccaca tggctgggag acagagtcac caccaccagc 1560  
 acccgcacct gggccctgcc cacctacaac aaccacctct acaagcagat atcaagtcag 1620  
 55  
 agcggggcta ccaacgacaa ccacttcttc ggctacagca cccctgggg ctattttgac 1680

# EP 1 310 571 A2

5	ttcaacagat tccactgccca cttctcacca cgtgactggc agcgactcat caacaacaac	1740
	tggggattcc ggcccagaaa gctgcggttc aagttgttca acatccaggt caaggaggtc	1800
	acgacgaacg acggcggttac gaccatcgct aataacctta ccagcacgat tcaggtcttc	1860
10	tcggactcgg agtaccact gccgtacgtc ctgggctctg cgcaccaggc ctgcctccct	1920
	ccgttccctg cggacgtgtt catgattcct cagtacggat atctgactct aaacaacggc	1980
	agtcagtctg tgggacgttc ctcttcttac tgcctggagt actttccttc tcagatgctg	2040
15	agaacgggca ataactttga attcagctac acctttgagg aagtgccttt ccacagcagc	2100
	tatgcgacaca gccagagcct ggaccggctg atgaatcccc tcacgacaca gtacctgtac	2160
	tacctggccc ggaccagag cactacgggg tccacaaggg agctgcagtt ccatcaggct	2220
20	gggccaaca ccatggccga gcaatcaaag aactggctgc cggaccctg ttatcggcag	2280
	cagagactgt caaaaaacat agacagcaac aacaacagta actttgcctg gaccggggcc	2340
	actaaatacc atctgaatgg tagaaattca ttaaccaacc cggcgctagc catggccacc	2400
25	aacaaggacg acgaggacca gttctttccc atcaacggag tgctggtttt tggcgaaacg	2460
	ggggctgccca acaagacaac gctggaaaac gtgctaataa ccagcgagga ggagatcaaa	2520
	accaccaatc ccgtggctac agaagaatac ggtgtggtct ccagcaacct gcaatcgtct	2580
30	acggccggac cccagacaca gactgtcaac agccaggggg ctctgcccgg catggtctgg	2640
	cagaaccggg acgtgtacct gcagggcccc atctgggcca aaattcctca cacggacggc	2700
	aactttcacc cgtctccct gatggcgga tttggactca aacaccggc tcctcaaatt	2760
35	ctcatcaaaa acacccggg acctgtaaat cctccagagg tgtttactcc tgccaagt	2820
	gcctcattta tcacgcagta cagcaccggc caggtcagcg tggagatoga gtgggaactg	2880
	cagaaagaaa acagcaaacg ctggaatcca gagattcagt acacctcaaa ttatgccaag	2940
40	tctaataatg tggaatttgc tgtcaacaac gaaggggttt atactgagcc tcgccccatt	3000
	ggcaccggtt acctcaccg taacctgtaa ttgcctgtta atcaataaac cggttaattc	3060
	gtttcagttg aactttggtc tctgcgaagg gcgaattc	3098
45	<210> 10	
	<211> 3098	
	<212> DNA	
	<213> new AAV serotype, clone 16.3	
	<400> 10	
50	gaattcgccc ttcgcagaga ccaaagttca actgaaacga atcaaccggg ttattgatta	60
	acaagtaatt acaggttacg ggtgaggtaa cgggtgccaa tggggcgagg ctcagtataa	120
	acccttcgt tggtgacagc aaattccaca ttattagact tggcataatt tgaggtgtac	180
55	tgaatctctg gattccagcg tttgctgttt tctttctgca gttccactc gatctccacg	240

# EP 1 310 571 A2

	ctgacctggc cgggtgctgta ctgctgtgata aatgaggcaa actaggcagg agtaaacacc	300
5	cctggaggat tagcaggtac cgggggtgttt ttgatgagaa tttgaggagg cgggtgtttg	360
	agtccaaatc cgcccatcag gggagacggg tgaaagtgc cgtccgtgtg aggaattttg	420
	gccagatgg gaccctgcag gtacacgtcc cggttctgcc agaccatgcc gggcagagcc	480
10	ccctggctgt tgacagtctg tgtctggggg cgggccgtag acgattgcag gttgctggag	540
	accacaccgt attcttctgt agccacggga ttggtgtgtt tgatctctc ctcgctggtc	600
	attagcacgt tttccagcgt tgtcttgttg gcagccccg ttttgccaaa aaccagcact	660
15	cgttgatgg gaaagaactg gccctcgtcg tccttgttgg tggccatggc tacgccggg	720
	ttggttaatg aatttctacc attcagatgg tatttagtgg ccccggtcca ggcaaagtta	780
	ctgttgttgt tgctgtctat gttttttgac agtctctgct gccgataaca gggctccgggc	840
20	agccagttct ttgattgctc ggccatggtg ttggggccag cctgatggaa ctgcagctcc	900
	cttgtggacc ccgtagtgtc ctgggtccgg gccaggtagt acaggtactg gtcgatgagg	960
	ggattcatca gccgggtccag gctctggctg tgccgatagc tgctgtggaa aggcacttcc	1020
25	tcaaaggtgt agctgaattc aaagtatttg cccgttctca gcactctgaga aggaaagtac	1080
	tccaggcagt agaaggagga acgtcccata gactgactgc cgttgtttag agtcagatat	1140
	ccgtactgag gaatcatgaa cacgtccgca gggaacggag ggaggcagcc ctggtgcgca	1200
30	gagccgagga cgtacggcag ttggtactcc gagtccgaga agacctgaat cgtgctggta	1260
	aggttattag cgatggtcgt aacgccgtcg ttctctgtga cctccttgac ctggatgttg	1320
	aacaacttga accgcagctt tctgggccgg aatccccagt tgttgttgat gagtctgtgc	1380
35	cagtacagtg gtgagaagtg gcagtggaat ctgttgaagt caaaatagcc ccaggggggtg	1440
	ctgtagccga agaagtgggt gtcgttggtg gccccgctct gacttgatat ctgctttag	1500
	agggtggtgt tgtaggtggg cagggcccg gtgcgggtgc tgggtggtgat gactctgtcg	1560
40	cccagccatg tggaatcgca atgccaattt ccggaggcat taccactcc gtcgggcct	1620
	tcgttattgt ctgccattgg tgcgccaccg cctgcagcca ttgtaccaga tcccagacct	1680
	gagggcgcgg cgggagggtc tccgagaggt tgggggtcgg gcactgactc tgagtcgcca	1740
45	gtctgcccga agttgagctt ctttttagcg ggctgctggc ctttcttgcc gatgcccggtg	1800
	gaggagtccg gggattctat gggctctctc tttccaggag ccgtcttagc gccttctca	1860
	accagaccga gaggttcgag aaccgcttc ttggcctgga agactgctcg cccgaggttg	1920
50	ccccaaaag acgtatcttc ttgaagacgc tcctgaaact cagcgtcggc gtggttgtac	1980
	ttgaggtagc ggttgtcccc ctgctcgagc tgcttgtcgt aggcctgtc gtgctcgagg	2040
	gccgcggcgt ctgcctcgtt gaccggtctc cccttgtcga gtccgttgaa gggctccagg	2100
55	tacttgtagc caggaagcac cagaccccg ccgtcgtcct gcttttgctg gttggccttg	2160

# EP 1 310 571 A2

	ggtttcgggg ctccagggtt caagtccac cactcgcgaa tgccctcaga gaggttggtcc	2220
5	tcgagccaat ctggaagata accatcggca gccatacctg gtttaagtca tttattgctc	2280
	agaaacacag tcatccagggt ccacgttgac cagatcgcag gccgagcaag caatctcggg	2340
	agcccccccc agcagatgat gaatggcaca gagtttccga tacgtcctct ttctgacgac	2400
10	cggttgagat tctgacacgc cggggaaaca ttctgaacag tctctgggtcc cgtgcgtgaa	2460
	gcaaagtgtt aaattctgat tcattctctc gcatgtcttg cagggaaaca gcatctgaag	2520
	catgcccgcg tgacgagaac atttgttttg gtacctgtcg gcaaagtcca ccggagctcc	2580
15	ttccgcgtct gacgtcgatg gatccgcgac tgaggggcag gcccgcttgg gctcgctttt	2640
	atccgcgtca tcggggggcg gcctcttggt ggctccacc tttctgacgt agaactcatg	2700
	cgccacctcg gtcacgtgat cctgcgcca gcggaagaac tctttgactt cctgctttgt	2760
20	caccttgcca aagtcttgct ccagacggcg ggtgagttca aatttgaaca tccggtcttg	2820
	taacggctgc tgggtgctga aggtggtgct gttcccgta atcacggcg acatgttggt	2880
	gttggaagtg acgatcacgg ggggtggatc gatctggcg gacgacttg acttttggtc	2940
25	cacgcgcacc ttgtgcccgc cgagaatggc cttggcgac tccacgacct tggccgtcat	3000
	cttgccctcc tcccaccaga tcaccatctt gtcgacgcaa tcgttgaagg gaaagttctc	3060
	attgggtccag ttgacgcagc cgtagaaagg gcgaattc	3098
30	<210> 11 <211> 3121 <212> DNA <213> new AAV serotype, clone 29.3	
35	<400> 11 gaattcgccc ttcgcagaga ccaaagttca actgaaacga atcaaccggt ttattgatta	60
	acaagcaatt acagattacg ggtgaggtaa cgggtgccga tggggcgagg ctcaataa	120
	gtgccatctg tgttaacagc aaagtccaca tttgtagatt tgtagtagtt ggaagtgtat	180
40	tgaatctctg gggtccagcg tttgctgttt tctttctgca gctccattc aatttccacg	240
	ctgacctgtc cgggtgctgta ctgcgtgatg aacgacgcca gcttagcttg actgaaggta	300
	gttgaggat ccgcgggaac aggtgtattc ttaatcagga tctgaggagg cgggtgtttc	360
45	agtccaaagc ccccatcag cggcgaggga tgaaagtttc cgtccgtgtg aggaatcttg	420
	gccagatag gacctgcag gtacacgtcc cggttctgcc agaccatgcc aggttaaggct	480
	ccttgactgt tgacggcccc tacaatagga cgggcgtttt gctgttgca gttatcggcc	540
50	accacgccgt actgttctgt ggccactggg ttggtggttt taatttcttc ctactggtt	600
	agcataacgc tgctatagtc cacgttgcc tttccagctc cctgtttccc aaacattaag	660
	actccgctgg acggaaaaaa tcgctcttcg tcgtccttgt gggttgccat agcgacaccg	720
55	ggatttacca gagagtctct gccattcaga tgatacttgg tggcaccggt ccaggcaaaag	780

# EP 1 310 571 A2

5	ttgctgttgt tattttgcga cagtgtcgtg gagacgcgtt gctgccggtg gcagggcccg	840
	ggtagccagt ttttggcctg agccgacatg ttattaggcc cggcctgaga aaatagcaac	900
	tgctgagttc ctgcggtacc tcccgtggac tgagtcggag acaggtagta caggtactgg	960
10	tcgatgaggg ggttcatcag ccggtccagg ctttggctgt gccgctagct gctgtgaaaa	1020
	ggcacgtcct caaactggta gctgaactca aagtgtgtgc ccgttctcag catttgagaa	1080
	ggaaagtact ccaggcagta gaaggaggaa cggcccacgg cctgactgcc attgttcaga	1140
15	gtcaggtagc cgtactgagg aatcatgaag acgtccggc ggaacggagg caggcagccc	1200
	tggcgcgcag agccgaggac gtacgggagc tggattccg agtccgtaaa gacctgaatc	1260
	gtgctggtaa ggttattggc gatggctctg gtgccttcat tctgcgtgac ctccctgacc	1320
20	tggatgttga agagcttgaa gttgagtctc ttgggcccga atccccagtt gttgttgatg	1380
	agtcgctgcc agtcacgtgg tgagaagtgg cagtggaaac tgttaaagtc aaaatacccc	1440
	cagggggtgc tgtagccgaa gtaggtgttg tcgttggtgc ttccctccga agtcccgttg	1500
25	gagatttgct tgtagaggtg gttgtttag gtggggaggg ccaggttcg ggtgctggtg	1560
	gtgatgactc tgtcgcccag ccatgtggaa tcgcaatgcc aatttctga ggaactacct	1620
	actccgtcgg cgccttcgtt attgtctgcc attggagcgc caccgcctgc agccattgta	1680
30	ccagatccca gaccagaggg gcctgcgggg ggttctccga ttggttgagg gtcgggact	1740
	gactctgagt cgccagtctg cccaaagtgt agtctctttt tcgcgggctg ctggcctttc	1800
	ttgccgatgc ccgtagtggg gtctggagaa cgctggggtg atggctctac cggctctctc	1860
35	tttccaggag ccgtcttagc gccttcctca accagaccga gaggttcgag aaccgccttc	1920
	ttggcctgga agactgctcg tccgaggttg ccccaaaaag acgtatcttc ttgcagacgc	1980
	tcctgaaact cggcgtcggc gtggttatac cgcaggtacg gattgtcacc cgctttgagc	2040
40	tgctggctgt aggccttgtc gtgctcgagg gccgctgcgt ccgcccgcgt gacgggctcc	2100
	cccttgctga gtccgttgaa ggggtccagg tacttgtagc caggaagcac cagaccccg	2160
	ccgtcgtcct gcttttgctg gttggctttg ggcttcgggg ctccagggtt cagcggccac	2220
45	cactcgcgaa tgccctcaga gaggttgctc tcgagccaat ctggaagata accatcggca	2280
	gccatacctg atctaaatca tttattgttc aaagatgcag tcatccaaat ccacattgac	2340
	cagatcgtag gcagtgcagg cgtctggcac ctttcccatg atatgatgaa ttagtcacag	2400
50	tttctgatac gcctttttga cgacagaaac ggggttgagat tctgacacgg gaaagcactc	2460
	taaacagtct ttctgtccgt gagtgaagca gatatttgaa ttctgattca ttctctcgca	2520
	ttgtctgcag ggaaacagca tcagattcat gccacgtga cgagaacatt tgttttggtg	2580
55	cctgtccgag tagttgatg aagcttcgac gtctgacgtc gatggctgag caactgactc	2640
	gcgcacccgt ttgggctcac ttatatctgc gtcactgggg gcgggtcttt tcttggtctc	2700

# EP 1 310 571 A2

5 accctttttg acgtagaatt catgctccac ctcaaccacg tgatcctttg cccaccggaa 2760  
 aaagtctttg acttcctgct tgggtgacctt cccaaagtca tgatccagac ggcgggtgag 2820  
 ttcaaatttg aacatccggt cttgcaacgg ctgctggtgt tcgaaggteg ttgagttccc 2880  
 10 gtcaatcacg gcgcacatgt tgggtgttggg ggtgacgac acgggagteg ggtctatctg 2940  
 ggccgaggac ttgcatttct ggtccacgcg caccttgctt cctccgagaa tggctttggc 3000  
 cgactccacg accttggcgg tcatcttccc ctctctccac cagatcacca tcttgtcgac 3060  
 15 acagtcgttg aagggaagt tctcattggt ccagttgacg cagccgtaga agggcgaatt 3120  
 c 3121  
  
 <210> 12  
 <211> 3121  
 20 <212> DNA  
 <213> new AAV serotype, clone 29.4  
  
 <400> 12  
 gaattcgccc ttctacggct gcgtcaactg gaccaatgag aactttccct tcaacgactg 60  
 25 tgtcgacaag atggtgatct ggtgggagga ggggaagatg accgccaagg tcgtggagtc 120  
 ggccaaagcc attctcggag gaagcaaggt gcgcgtggac cagaaatgca agtcctcggc 180  
 ccagatagac ccgactcccg tgatcgtcac ctccaacacc aacatgtgcg ccgtgattga 240  
 30 cgggaactca acgacctcg aacaccagca gccgttgcaa gaccggatgt tcaaatttga 300  
 actcaccgcg cgtctggatc atgactttgg gaaggtcacc aagcaggaag tcaaagactt 360  
 tttccgggtg gcaaaggatc acgtggttga ggtggagcac gaattctacg tcaaaaaggg 420  
 35 tggagccaag aaaagacccg cccccagtga cgcagatata agtgagcca aacgggtgcg 480  
 cgagtcagtt gcgcagccat cgacgtcaga cgcggaagct tcgatcaact acgcagacag 540  
 gtacaaaaac aaatgttctc gtcacgcggg catgaatctg atgctgtttc cctgcagaca 600  
 40 atgcgagaga atgaatcaga attcaaatat ctgcttcact cacggacaga aagactgttt 660  
 agagtgcctt cccgtgtcag aatctcaacc cgtttctgtc gtcaaaaagg cgtatcagaa 720  
 actgtgctac attcatcata tcatgggaaa ggtgccagac gcttgcaact cctgcgatct 780  
 45 ggtcgatgtg gatttggatg actgcatctt tgaacaataa atgatttaaa tcaggatatg 840  
 ctgccgatgg ttatcttcca gattggctcg aggacaacct ctctgagggc attcgcgagt 900  
 ggtgggcgct gaaacctgga gccccgaag ccaaagccaa ccagcaaaag caggacggcg 960  
 50 gccggggtct ggtgcttctt ggctacaagt acctcggacc cttcaacgga ctcgacaagg 1020  
 gggagcccggt caacgcggcg gacgcagcgg ccctcgagca cgacaaggcc tacgaccagc 1080  
 agctcaaagc gggtgacaat ccgtacctgc ggtataacca cgccgacgcc gattttcagg 1140  
 55 agcgtctgca agaagatacg tcttttgggg gcaacctcgg gcgagcagtc ttccaggcca 1200

# EP 1 310 571 A2

	agaagcgggt tctcgaacct ctcggtctgg ttgaggaagg cgctaagacg gctcctggaa	1260
5	agaagagacc ggtagagcca tcaccccagc gttctccaga ctctctacg ggcacgcgca	1320
	agaaaggcca gcagcccgcg aaaaagagac tcaactttgg gcagactggc gactcagagt	1380
	cagtgcccgga ccctcaacca atcggagaaac cccccgcagg cccctctggg ctgggatctg	1440
10	gtacaatggc tgcaggcggg ggcgctccaa tggcagacaa taacgaaggc gccgacggag	1500
	tgggtagttc ctacaggaaat tggcattgcg attccacatg gctgggcgac tgagtcatca	1560
	ccaccagcac ccgaacctgg gccctcccca cctacaacaa ccacctctac aagcaaatct	1620
15	ccaacgggac ttcgggagga agcaccaacg acaacaccta cttcggctac agcaccccct	1680
	gggggtatatt tgactttaac agattccact gccactttct accacgtgac tggcagcgac	1740
	tcatcaacaa caactgggga ttccggccca agagactcaa cttcaagctc ttcaacatcc	1800
20	agggtcaagga ggtcacgcag aatgaaggca ccaagaccat cgccaataac cttaccagca	1860
	cgattcaggt ctttacggac tcggaatacc agctcccgta cgtcctcggc tctgcgcacc	1920
	agggctgcct gcctccgttc ccggcggacg tcttcatgat tcctcagtac gggtagctga	1980
25	ctctgaacaa tggcagtcag gccgtgggac gttcctcctt ctactgcctg gagtactttc	2040
	cttctcaaat gctgagaacg ggcaacaact ttgagttcag ctaccagttt gaggacgtgc	2100
	cttttcacag cagctacgag cacagccaaa gcctggaccg gctgatgaac cccctcatcg	2160
30	accagtacct gtactacctg tctcggactc agtccacggg aggtaccgca ggaactcagc	2220
	agttgctatt ttctcaggcc gggcctaata acatgtcggc tcaggccaaa aactggctac	2280
	ccgggcccctg ctaccggcag taacgcgtct ccacgacact gtcgcaaaat aacaacagca	2340
35	actttgtctg gaccggtgcc accaagtatc atctgaatgg cagagactct ctggtagatc	2400
	ccggtgtcgc tatggcaacc cacaaggacg acgaagagcg attttttccg tccagcggag	2460
	tcataatgtt tgggaaacag ggagctggaa aagacaacgt ggactatagc agcgtcatgc	2520
40	taaccagtga ggaagaaatt aaaaccacca acccagtggc cacagaacag tacggcgtgg	2580
	tggccgataa cctgcaacag caaaacgccg ctctatttgt aggggccgtc aacagtcaag	2640
	gagccttacc tggcatggc tggcagaacc gggacgtgta cctgcagggt cctacctggg	2700
45	ccaagattcc tcacacggac ggaaactttc atccctcgcc gctgatggga ggctttggac	2760
	tgaaacaccc gcctcctcag atcctgatta agaatacacc tgttcccgcg gatcctccaa	2820
	ctaccttcag tcaagctaag ctggcgtcgt tcatcacgca gtacagcacc ggacagggtca	2880
50	gcgtggaaat tgaatgggag ctgcaggaag aaaacagcaa acgctggaac ccagagattc	2940
	aatacacttc caactactac aaatctacaa atgtggactt tgctgttaac acagatggca	3000
	cttattctga gcctcgcccc atcggcaccg gttacctcac ccgtaatctg taattgcttg	3060
55	ttaatcaata aaccggttga ttcgtttcag ttgaactttg gtctctgcga agggcgaatt	3120

# EP 1 310 571 A2

	c	3121
5	<210> 13 <211> 3121 <212> DNA <213> new AAV serotype, clone 29.5	
10	<400> 13 gaattcgccc ttcgcgagac caaagttcaa ctgaaacgaa tcaaccggtt tattgattaa	60
	caagcaatta cagattacgg gtgaggtaac ggggtccgat ggggagagc tcagaataag	120
	tgccatctgt gttacagca aagtcacat ttgtagattt gtagtagttg gaagtgtatt	180
15	gaatctctgg gttccagcgt ttgctgtttt ctttctgcag ctccattca atttcacgc	240
	tgacctgtcc ggtgctgtac tgcgtgatga acgacgccag cttagcttga ctgaaggtag	300
	ttggaggatc cgcgggaaca ggtgtattct taatcaggat ctgaggaggc ggggtgtttca	360
20	gtccaaagcc tcccatcagc ggcgagggat gaaagtttcc gtccgtgtga ggaatcttgg	420
	cccagatagg accctgcagg tacacgtccc ggttctgcca gaccatgcca ggtaaggctc	480
	cttgactgtt gacggcccct acaataggag cggcgttttg ctgttgacag ttatcggcca	540
25	ccacgccgta ctgttctgtg gccactgggt tgggtgtttt aatttcttcc tctactggtta	600
	gcataacgct gctatagtc acgttgtctt ttccagctcc ctgtttccca aacattaaga	660
	ctccgctgga cggaaaaaat cgctcttctg cgtccttgtg ggttgccata gcgacaccgg	720
30	gatttaccag agagtctctg ccattcagat gatacttggg ggcaccggtc caggcaaagt	780
	tgctgttggtc attttgcgac agtgctcgtg agacgcgttg ctgccggtag cagggcccg	840
	gtagccagtt tttggcctga gccgacatgt tattaggccc ggcctgagaa aatagcaact	900
35	gctgagttcc tgcggtacct cccgtggact gagtccgaga caggtagtac aggtactgg	960
	cgatgagggg gttcatcagc cggtcaggc tttggctgtg cgcgtagctg ctgtgaaaag	1020
	gcacgtctc aaactggtag ctgaactcaa agttgttgcc cgttctcagc atttgagaag	1080
40	gaaagtactc caggcagtag aaggaggaac ggccacggc ctgactgcca ttgttcagag	1140
	tcaggtagcc gtactgagga atcatgaaga cgtccgcccga gaacgagggc aggcagccct	1200
	gggtgcgcaga gccgaggacg tacgggagct ggtattccga gtccgtaaag acctgaatcg	1260
45	tgctggtgaa gttattggcg atggtcttgg tgccttcatt ctgctgacc tccttgacct	1320
	ggatgttgaa gagcttgaag ttgaggctct tgggcccga tccccagttg ttgttgatga	1380
	gtcgtgcca gtcacgtggt gagaagtggc agtggaaatct gttaaagtca aaatacccc	1440
50	aggggggtgct gtagccgaag taggtgttgt cgttggtgct tcctcccgaa gtcccgttgg	1500
	agatttgctt gtagagggtg ttgttgtagg tggggagggc ccaggttcgg gtgctggtg	1560
	tgatgactcc gtcgcccagc catgtggaat cgcaatgcca atttctgag gaactacca	1620
55	ctccgtcggc gccttcgtta ttgtctgcca ttggagcgcc accgcctgca gccattgtac	1680

# EP 1 310 571 A2

	cagatccag accagagggg cctgcggggg gttctccgat tggttgaggg tcgggcactg	1740
5	actctgagtc gccagtcctgc ccaaagttga gtctcttttt cgcgggctgc tggcctttct	1800
	tgccgatgcc cgtagaggag tctggagaac gctggggtga tggctctacc ggtctcttct	1860
	ttccaggagc cgtcttagcg ccttcctcaa ccagaccgag aggttcgaga acccgcttct	1920
10	tggcctggaa gactgctcgc ccgaggttgc ccccaaaaga cgtatcttct tgcagacgct	1980
	cctgaaactc ggcgtcggcg tggttatacc gcaggtacgg attgtcaccg gctttgagct	2040
	gctggtcgta ggccttgctg tgctcgaggg ccgctgcgct cgccgcgttg acgggctccc	2100
15	ccttgctcag tccgttgaag ggtccgaggt acttgtagcc aggaagcacc agaccccggc	2160
	cgctcgtcctg cttttgctgg ttggctttgg gcttcggggc tccaggtttc agcgcccacc	2220
	actcgcaat gccctcagag aggttgctct cgagccaatc tggaagataa ccatcggcag	2280
20	ccatacctga tttaaatcat ttattgttca aagatgcagt catccaaatc cacattgacc	2340
	agatcgcagg cagtgaagc gtctggcacc tttcccatga tatgatgaat gtagcacagt	2400
	ttctgatacg cttttttgac gacagaaacg ggttgagatt ctgacacggg aaagcactct	2460
25	aaacagtcct tctgtccgtg agtgaagcag atatttgaat tctgattcat tctctcgc	2520
	tgtctgcagg gaaacagcat cagattcatg cccacgtgac gagaacattt gttttggtac	2580
	ctgtctgcgt agttgatcga agcttcgcgc tctgacgtcg atggctgcgc aactgactcg	2640
30	cgcacccggt tgggctcact tatactctcg tcactggggg cgggtctttt cttggctcca	2700
	ccctttttga cgtagaatc atgctccacc tcaaccacgt gatcctttgc ccaccgaaa	2760
	aagtctttga cttcctgctt ggtgaccttc ccaaagtcac gatccagacg gcgggtgagt	2820
35	tcaaatttga acatccggtc ttgcaacggc tgctggtgtt cgaaggtcgt tgagttcccg	2880
	tcaatcacgg cgcacatgtt ggtgttgagg gtgacgatca cgggagtcgg gtctatctgg	2940
	gccgaggact tgcatttctg gtccacgcgc accttgcttc ctccgagaat ggctttggcc	3000
40	gactccacga cttggcggt catcttcccc tcctcccacc agatcaccat cttgtcgaca	3060
	cagtcgttga agggaaagtt ctcatgtgtc cagttgacgc agccgtagaa agggcgaatt	3120
	c	3121
45		
	<210> 14	
	<211> 3131	
	<212> DNA	
	<213> new AAV serotype, clone 1-3	
50		
	<400> 14	
	gcggccgcga attcgccctt ggctgcgtca actggaccaa tgagaacttt cccttcaatg	60
	attgcgtcga caagatggtg atctggtggg aggagggcaa gatgacggcc aaggtcgtgg	120
55	agtccgccaa ggccattctc ggcggcagca aggtgcgcgt ggaccaaag tgcaagtcgt	180

# EP 1 310 571 A2

	ccgcccagat cgaccccacc cccgtgatcg tcacctccaa caccaacatg tgcgccgtga	240
5	ttgacgggaa cagcaccacc ttcgagcacc agcagcctct ccaggaccgg atgtttaagt	300
	tcgaactcac ccgccgtctg gaggcacgact ttggcaaggt gacaaagcag gaagtcaaag	360
	agttcttccg ctgggccagt gatcacgtga ccgagggtggc gcatgagttt tacgtcagaa	420
10	agggcggagc cagcaaaaga cccgcccccg atgacgcgga taaaagcgag cccaagcggg	480
	cctgccctc agtcgcggat ccatcgacgt cagacgcgga aggagctccg gtggactttg	540
	ccgacaggta ccaaaacaaa tgtttctcgtc acgcgggcat gcttcagatg ctgtttccct	600
15	gcaaaacgtg cgagagaatg aatcggaatt tcaacatttg cttcacacac ggggtcagag	660
	actgctcaga gtgtttcccc ggcgtgtcag aatctcaacc ggtcgtcaga aagaggacgt	720
	atcggaact ccgtgcgatt catcatctgc tggggcgggc tcccagagatt gcttgctcgg	780
20	cctgcgatct ggtcaacgtg gacctggatg actgtgtttc tgagcaataa atgacttaaa	840
	ccaggatatg ctgccgatg ttatcttcca gattggctcg aggacaacct ctctgagggc	900
	attcgcgagt ggtgggcgct gaaacctgga gccccgaagc ccaaagccaa ccagcaaaag	960
25	caggacgacg gccggggtct ggtgcttctt ggctacaagt acctcggacc cttcaacgga	1020
	ctcgacaagg gggagcccgt caacgcggcg gacgcagcgg ccctcgagca cgacaaggct	1080
	tacgaccagc agctgcaggc gggtgacaat ccgtacctgc ggtataacca cgccgacgcc	1140
30	gagtttcagg agcgtctgca agaagatacg tcttttgggg gcaacctcgg gcgagcagtc	1200
	ttccaggcca agaagcgggt tctcgaaact ctcggtctgg ttgaggaagg cgctaagacg	1260
	gctcctggaa agaagagacc ggtagagcca tcacccagc gttctccaga ctctctacg	1320
35	ggcatcggca agaaaggcca acagcccgc agaaaaagac tcaattttgg tcagactggc	1380
	gactcagagt cagttccaga ccctcaacct ctcgagagac ctccagcagc gccctctggt	1440
	gtgggaccta atacaatggc tgcaggcggg ggcgcaccaa tggcagacaa taacgaaggc	1500
40	gccgacggag tgggtagtct ctcgggaaat tggcattgag attccacatg gctgggcgac	1560
	agagtcatca ccaccagcac ccgaacctgg gccctgcccc cctacaacaa ccacctctac	1620
	aagcaaatct ccaacgggac atcgggagga gccaccaacg acaacaccta cttcggctac	1680
45	agcaccacct gggggtatct tgactttaac agattccact gccacctttc accacgtgac	1740
	tggcagcgac tcatcaacaa caactgggga ttccgaccca agagactcag cttcaagctc	1800
	ttcaacatcc aggtcaagga ggtcacgcag aatgaaggca ccaagaccat cgccaataac	1860
50	ctcaccagca ccatccagggt gtttacggac tcggagtacc agctgccgta cgttctcggc	1920
	tctgtccacc agggctgcct gcctccgttc ccggcggacg tgttcatgat tcccagtac	1980
	ggctacctaa cactcaacaa cggtagtcag gccgtgggac gctcctcctt ctactgcctg	2040
55	gaatactttc cttcgcagat gctgagaacc ggcaacaact tccagtttac ttacaccttc	2100

# EP 1 310 571 A2

	gaggacgtgc ctttccacag cagctacgcc cacagctaga gcttggaccg gctgatgaat	2160
5	cctctgattg accagtacct gtactacttg tctcggactc aaacaacagg aggcacggca	2220
	aatacgcaga ctctgggctt cagccaaggt gggcctaata caatggccaa tcaggcaaag	2280
	aactggctgc caggaccctg ttaccgcca caacgcgtct caacgacaac cgggcaaaac	2340
10	aacaatagca actttgcctg gactgctggg accaaatacc atctgaatgg aagaaattca	2400
	ttggctaatac ctggcatcgc tatggcaaca cacaagacg acgaggagcg tttttttccc	2460
	agtaacggga tcctgatttt tggcaaacaa aatgctgcc gagacaatgc ggattacagc	2520
15	gatgtcatgc tcaccagcga ggaagaaatc aaaaccacta accctgtggc tacagaggaa	2580
	tacggtatcg tggcagataa cttgcagcag caaaacacgg ctctcaaat tggaactgtc	2640
	aacagccagg gggccttacc cggtatggtc tggcagaacc gggacgtgta cctgcagggt	2700
20	cccatctggg ccaagattcc tcacacggac ggcaacttcc acccgtctcc gctgatgggc	2760
	ggcttttggc tgaaacatcc tccgcctcag atcctgatca agaacacgcc tgtacctgcy	2820
	gacccctcca ccaccttcaa ccagtcaaag ctgaactctt tcattcacgca atacagcacc	2880
25	ggacaggtca gcgtggaaat tgaatgggag ctgcagaagg aaaacagcaa gcgctggaac	2940
	cccgagatcc agtacacctc caactactac aaatctataa gtgtggactt tgctgttaat	3000
	acagaaggcg tgtactctga accccgcccc attggcaccg gttacctcac ccgtaatctg	3060
30	taattgcctg ttaatcaata aaccggttga ttcgtttcag ttgaactttg gtctctgcga	3120
	agggcgaatt c	3131
35	<210> 15 <211> 3127 <212> DNA <213> new AAV serotype, clone 13-3b	
40	<400> 15 gcggccgcga attcgccctt cgcagagacc aaagttcaac tgaaacgaat caaccggttt	60
	attgattaac atgcaattac agattacggg tgaggtaacg agtgccaata gggcgaggct	120
	cagagtaaac accctggctg tcaacggcaa agtccacacc agtctgcttt tcaaagttag	180
	aggtgtactg aatctccggg tcccagcgct tgctgttttc cttctgcagc tcccactoga	240
45	tttccacgct gacttgtccg gtgctgtact gtgtgatgaa cgaagcaaac ttggcaggag	300
	taaacacctc cggaggatta gcgggaacgg gagtgttctt gatcaggatc tgaggaggcg	360
	gatgtttaag tccaaagccg cccatcaaag gagacgggtg aaagltgcca tccgtgtgag	420
50	gaatcttggc ccagatggga ccctgcaggc acacgtcccg gttctgccag accatgccag	480
	gtaaggctcc ctggttgttg acaacttgtg tctgggctgc agtattagcc gcttgtaagt	540
	tgctgtgac tatcccgat tcttccgtgg ctacaggatt agtaggacga atttcttctt	600
55	catttgtcat taacacattt tccaatgtag ttttgttagt tgctccagtt tttccaaaaa	660

EP 1 310 571 A2

5	tcaggactcc gctggatggg aaaaagcggg cctcgtcgtc cttgtgagtt gccatggcga	720
	cgccgggatt aaccaacgag tttctgccgt tcagggtgata tttggtggca ccagtccaag	780
	caaagttagct gttgttgttt tgatccagcg ttttggagac ctttgtttgc cggaagcagg	840
10	gtccaggtaa ccaattcttg gcttgttcgg ccatagttga agggccgccc tggtaaaact	900
	gcagttcccc attgccagct gtgcctcctg ggtcactctg tgttctggcc aggtagtaca	960
	agtactggtc gatgagggga ttcacagacc ggtccaggct ctggctgtgt gcgtagctgc	1020
15	tgtggaaaag cagtcctcg aagctgtagc tgaactcaaa gttgttggcc gttctcagca	1080
	tctgagaggg gaagtactcc aggcagtaga aggaggaacg tcccacagac tgactgccat	1140
	tgttgagagt caggtagccg tactgaggaa tcatgaagac gtccgccggg aacggaggca	1200
20	ggcagccctg gtgcgcagag ccgaggacgt acggcagctg gtattccgag tccgagaata	1260
	cctgaatcgt gctggtaaag ttattagcga tggctgtaac gccgtcatte gtcgtgacct	1320
	ccttgacctg gatgttgaag agcttgaacc gcagcttctt gggccggaat ccccagttgt	1380
25	tgttgatgag tcgctgccag tcacgtgggtg agaagtggca gtggaatctg ttaaagtcaa	1440
	aatacccca gggggtgctg tagccgaagt aggtgtgtgc gttggtacta cctgcagttt	1500
	cactggagat ttgctcgtag aggtggttgt tgtaggtggg cagggcccgag gttcgggtgc	1560
30	tgggtgtaat gactctgtcg ccagccatg tggaatcgca atgccaattt cctgaggcat	1620
	taccactcc gtcggcacct tcgttattgt ctgccattgg tgcgccaccg cctgcagcca	1680
	ctgtaccaga tcccacacta gagggcgctg ctggaggttc tccgagaggt tgagggtcgg	1740
35	ggactgactc tgagtcgcca gtctgaccga aattgagtct ctttctggcg ggctgctggc	1800
	ccttcttgcc gatgcccggt gaggagtcgg gggaacgctg aggtgacggc tctaccggtc	1860
	tcttctttgc aggagccgtc ttagcgctt cctcaaccag accgagaggt tcgagaacct	1920
40	gcttcttgcc ctggaagact gctcgcccga ggttgccccc aaatgacgta tcttcttgca	1980
	gacgctcctg aaactcggcg tcggcgtggt tataccgcag gtacgggttg tcaccgcat	2040
	tgagctgctg gtcgtaggcc ttgtcgtgct cgagggccgc tgcgtccgcc gcgttgacgg	2100
45	gctccccctt gtcgagtcgg ttgaagggtc cgaggtactt gtagccagga agcaccagac	2160
	cccggccggt gtccctgctt tgctggttgg ctttgggttt cggggctcca ggtttcaggt	2220
	cccaccactc gcgaatgccc tcagagaggt tgcctcagag ccaatctgga agataaccat	2280
50	cggcagccat acctgattta aatcatttat tgttcaaaga tgcagtcac caaatccaca	2340
	ttgaccagat cgcaggcagt gcaagcgtct ggcaccttcc ccatgatatg atgaatgtag	2400
	cacagtttct gatacgcctt tttgacgaca gaaacgggtt tagattctga cacgggaaag	2460
55	cactctaaac agtctttctg tccgtgagtg aagcagatat ttgaattctg attcattctc	2520
	tcgcattgtc tgcagggaaa cagcatcaga ttcatgccc cgtgacgaga acatttgttt	2580

# EP 1 310 571 A2

5	tggtacctgt ctgcgtagtt gatcgaagct tccgcgtctg acgtcgatgg ctgcgcaact	2640
	gactcgcgca cccgtttggg ctcacttata tctgcgtcac tggggggcggg tctttttcttg	2700
	gctccaccct ttttgacgta gaattcatgc tccacctcaa ccacgtaatc ctttgcccac	2760
10	cggaaaaagt ctttgacttc ctgcttgggtg accttcccaa agtcatgac cagacggcgg	2820
	gtgagttcaa atttgaacat ccggctcttg aacggctgct ggtgttcgaa ggtcgttgag	2880
	ttcccgtcga tcacggcgca catgttgggtg ttggagatga cgatcgccgg agtcgggtct	2940
15	atctgggccc aggacttgca tttctggtcc acgcgcacct tgcttcctcc gagaatggct	3000
	ttggccgact ccacgacctt ggcggtcac ttcccctcct cccaccagat caccatcttg	3060
	tcgacacagt cgttgaaggg aaagtctctca ttggtccagt tgacgcagcc gtagaaagg	3120
20	cgaattc	3127
	<210> 16	
	<211> 3106	
	<212> DNA	
	<213> new AAV serotype, clone 24-1	
25	<400> 16	
	gcggccgcga attcgccctt cgcagagacc aaagtccaac tgaaacgaat caaccggttt	60
	attgattaac aagtaattac aggttacggg tgaggtaacg ggtgccaatg gggcgaggct	120
30	cagtataaac cccttcgttg ttgacagcaa attccacatt attagacttg gcataatttg	180
	aggtgtactg aatctctgga ttccagcggt tgctgttttc tttctgcagt tcccactcga	240
	tctccacgct gacctggccg gtgctgtact gcgtgataaa tgaggcaaac ttggcaggag	300
35	taaacacctc tggaggatta gcaggtaccg ggggtgtttt gatgagaatt tgaggaggcg	360
	ggtgtttgag tccaaatccg cccatcaggg gagacgggtg aaagttgccg tccgtgtgag	420
	gaattttggc ccagatggga ccctgcaggc acacgtcccg gttctgccag accatgccgg	480
40	gcagagcccc ctggctgttg acagtctgtg tctgggggtc gcccgtagac gattgcaggt	540
	tgctggagac cacaccgtat tcttctgtag ccacgggatt ggtggttttg atctcctcct	600
	cgctgggtcat tagcacgttt tccagcggtg tcttgttggc agccccggt ttgccaaaaa	660
45	ccagcactcc gttgatggga aagaactggt cctcgctcgtc cttgttgggt gccatggcta	720
	cgcccggtt ggttaatgaa tttctacat tcagatggta tttagtggcc ccggtccagg	780
	caaagttact gttgttgttg ctgtctatgt tttttgacag tctctgctgc cgataacagg	840
50	gtccgggag ccagttcttt gattgctcgg ccatggtgtt gggcccagcc tgatggaact	900
	gcagctccct tgtggacccc gtagtgctct gggtcgggc caggtagtac aggtactggt	960
	cgatgagggg attcatcagc cggctcaggc tctggctgtg cacatagctg ctgtggaaag	1020
55	gcacttcttc aaagggtgtag ctgaattcaa agttattgcc cgttctcagc atctgagaag	1080

EP 1 310 571 A2

	gaaagtactc caggcagtag aaggaggaac gtccacaga ctgactgccg ttgttttagag	1140
5	tcagatatcc gtactgagga atcatgaaca cgtccgcagg gaacggaggg aggcagccct	1200
	ggtgcgcaga gccgaggacg tacggcagtt ggtactccga gtccgagaag acctgaatcg	1260
	tgctggttaag gttattagcg atggtcgtaa cgccgtcggt cgtcgtgacc tccttgacct	1320
10	ggatgttgaa caacttgaa cgcagctttc tgggcccggaa tccccagttg ttgttgatga	1380
	gtcgtgcca gtcacgtggt gagaagtggc agtggaatct gttgaagtca aaatagcccc	1440
	agggggtgct gtagctgaag aagtgggtgt cgttggtagc cccgctctga cttgatatct	1500
15	gctttagtag gttggtgttg tagtgggca gggcccagggt gcgggtgctg gtggtgatga	1560
	ctctgtcgcc cagccatgtg gaatcgcaat gccaatctcc ggaggcatta cccactccgt	1620
	cggcgccctc gttattgtct gccattggtg cgccaccgcc tgcagccatt gtaccagatc	1680
20	ccagacctga gggcgccggc ggaggttctc cgagagggtg ggggtcgggc actgactctg	1740
	agtcgccagt ctgccccaaag ttgagcttct ttttagcggg ctgctggcct ttcttgccga	1800
	tgcccggtga ggagtcgggg gattctatgg gtctcttctt tccaggagcc gtcttagcga	1860
25	cttctcaac cagaccgaga ggttcgagaa cccgcttctt ggcttgaag actgctcgcc	1920
	cagaggttgc cccaaaagac gtatcttctt gaagacgctc ctgaaactcg gcgtcggcgt	1980
	ggttgactt gaggtacggg ttgtccccct gctcgagctg cttgtcgtag gccttgctgt	2040
30	gctcgagggc cgcggcgtct gcctcgttga ccggtctctc cttgtcgagt ccgttgaagg	2100
	gtctgaggta cttgtagcca ggaagcacca gaccccgcc gtctcctgc ttttctggt	2160
	tggccttggg tttcggggct ccaggtttca agtccacca ctgcgcaatg ccctcagaga	2220
35	ggttgctctc gagccaatct ggaagataac catcggcagc catacctggt ttaagtcat	2280
	tattgctcag aaacacagtc atccaggctc acgttgacca gatcgcaggc cgagcaagca	2340
	atctcgggag cccgccccag cagatgatga atggcacaga gtttccgata cgtcctcttt	2400
40	ctgacgaccg gttgagattc tgacacgccg gggaacatt ctgaacagtc tctggtcccg	2460
	tgctggaagc aaatgttgaa attctgattc actctctcgc atgtcttgca gggaaacagc	2520
	atctgaagca tgcccgcgtg acgagaacat ttgttttggg acctgtcggc aaagtccacc	2580
45	ggagctcctt ccgcgtctga cgtcgatgga ttccgcgactg aggggcaggc ccgcttgggc	2640
	tcgcttttat ccgcgtcatc gggggcgggt ctcttggttg cccaccctt tctgacgtag	2700
	aacctatgcg ccacctcggt cacgtgatcc tgcgccagc ggaagaacct tttgacttcc	2760
50	tgctttgtca ccttgccaaa gttatgctcc agacggcggg tgggttcaa tttgaacatc	2820
	cggctctgca acggtctgtg gtgctcgaag gtggcgctgt tcccgtaac cagggcgac	2880
	atgttggtgt tggaggtgac ggtcacgggg gtggggtcga tctggcgga cgacttgac	2940
55	ttttggtcca cgcgcacctt gctgccgccg agaattggcct tggcggactc cacgaccttg	3000

# EP 1 310 571 A2

gccgtcatct tgcctcctc ccaccagatc accatcttgt cggcgcaatc gttgaaggga 3060

5 aagttctcat tgggtccagtt gacgcagccg tagaaagggc gaattc 3106

<210> 17  
<211> 3102  
<212> DNA  
10 <213> new AAV serotype, clone 27-3

<400> 17  
gcggccgcga attcgccctt cgcagagacc aaagttcaac tgaaacgaat caaccggttt 60

attgattaac aagtaattac aggttacggg tgaggtaacg ggtgccaatg gggcgaggct 120

15 cagtataaac cccttcgttg ttgacagcaa attccacatt attagacttg gcataatttg 180

agggtgactg aatctctgga ttccagcgtt tgctgttttc tttctgcagt tccactcga 240

tctccacgct gacctggccg gtgctgtact gcgtgataaa tgaggcaaac ttggcaggag 300

20 taaacacctc tggaggatta gcaggtagcg ggggtgtttt gatgagaatt tgaggaggcg 360

ggtgttttag tccaaatccg cccatcaggg gagacgggtg aaagttgccg tccgtgtgag 420

gaatttcggc ccagatggga ccctgcaggt acacgtcccg gttctgccag accatgccgg 480

25 gcagagcccc ctggctgttg acagtctgtg tccggggtcc ggccgtagac gattgcaggt 540

tgctggagac cacaccgtat tcttctgtag ccacgggatt ggtgggtttg atctcctcct 600

cgctggatcat tagcacgttt tccagcgttg tcttgttgcc agcccccggt ttgccaaaaa 660

30 ccagcactcc gttgatggga aggaactggg cctcgtcgtc cttgttgggt gccatggcta 720

cgccccgggt gggttaatgaa tttctaccat tcagatggta tttagtggcc ccggtccagg 780

caaagttact gttgttgttg ctgtctatgt tttttgacag tctctgctgc cgataacagg 840

35 gtccgggcag ccagttcttt gattgctcgg ccacgggtgtt gggcccagcc tgatggaact 900

gcagctccct tgtggacccc gtagtgctct ggggtccggc caggtagtac aggtactggt 960

cgatgagggg attcatcagc cgggtccaggc tctggctgtg cgcatactgt ctgtggaaag 1020

40 gcacttcctc aaaggtgtag ctgaattcaa agttattgcc cgttctcagc atctgagaag 1080

gaaagtactc caggcagcag aaggaggaac gtcccacaga ctgactgccg ttgttttagag 1140

tcagatatcc gtactgagga atcatgaaca cgtccgcagg gaacggaggg aggcagccct 1200

45 ggtgcgcaga gccgaggacg tacggcagtt ggtactccga gtccgagaag acctgaatcg 1260

tgctggtaag gttattagcg atggctgtaa cgccgtcgtt cgtcgtgacc tccttgacct 1320

ggatgttgaa caacttgaac cgcagctttc tgggccggaa tccccagttg ttgttgatga 1380

50 gtcgctgccg gtcacgtggg gagaagtggc agtggaatct gttgaagtca aaatagcccc 1440

aggggggtgct gtagccgaag aagtggttgt cgttggtagc cccgctctga cttgatatct 1500

gcttgtagag gtggttgttg taggtgggca gggcccagggt gcgggtgctg gtggtgatga 1560

55 ctctgtcgcg cagccatgtg gaatcgcaat gccaatctcc ggaggcatta cccactccgt 1620

# EP 1 310 571 A2

5 cgggcgcttc gttattgtct gccattgggtg cgccaccgcc tgcagccatt gtaccagatc 1680  
 ccagacctga gggcgcggcg ggaggttctc cgagaggttg ggggtcgggc actgactctg 1740  
 agtcgccagt ctgccc aaag ttgagcttct ttttagcggg ctgctggcct ttcttgccga 1800  
 10 tgcccggtga ggagtcgggg gattctatgg gtctcttctt tccggaagcc gtcttagcgc 1860  
 ctteectaac cagaccgaga ggttcgagaa cccgcttctt ggcttgaag actgctcgcc 1920  
 cgaggttgcc cccaaaagac gtatcttctt gaagacgctc ctgaaactcg gcgtcggcgt 1980  
 ggttgtactt gaggtacggg ttgtccccct gctcgagctg cttgtcgtag gccttgctgt 2040  
 15 gctcgagggc cgcggcgtct gcctcgttga cgggctctcc cttgtcgtag ccgttgaagg 2100  
 gtccgaggtg cttgttagcca ggaagcacca gaccccgcc gtcgtcctgc ttttgctggt 2160  
 tggctttggg tttcggggct ccagggttca agtcccacca ctgcgcaatg cctcagaga 2220  
 20 ggttgtcctc gagccaatct ggaagataac catcggcagc catacctggt ttaagtcatt 2280  
 tattgctcag aaacacagtc atccaggctc acgttgacca gatcgagcgc cgagcaagca 2340  
 atctcgggag cccgccccag cagatgatga atggcacaga gtttccgata cgtcctcttt 2400  
 25 ctgacgaccg gttgagattc tgacacgcgc gggaaacatt ctgaacagtc tctggtcccg 2460  
 tgcgtgaagc aaatgttgaa attctgattc attctctcgc atgtcttgca gggaaacagc 2520  
 atctgaagca tgcccgctg acgagaacat ttgttttggg acctgtcggc aaagtccacc 2580  
 30 ggagctcctt ccgctgtctga cgtcgatgga tccgcgactg aggggcaagc ccgcttgggc 2640  
 tcgcttttat ccgctcatc gggggcgggg ctcttgttgg ctccaccctt tctgacgtag 2700  
 aactcatgcg ccacctcggg cagctgatcc tgcgcccagc ggaagaactc tttgacttcc 2760  
 35 tgctttgtca ccttgccaaa gtcagctc agacggcggg tgagttcaaa tttgaacatc 2820  
 cggctcttga acggctgctg gtgctcgaag gtggtgctgt tcccgtaac cagggcgac 2880  
 atgttggtgt tggaagtga gatcacgggg gtgggatcga tctgggcgga cgacttgac 2940  
 40 ttttggtcca cgcgacactt gctgccgcgc agaattggcct tggcggaact cagacacttg 3000  
 gccgtcatct tgccctctc ccaccagatc accatcttgt cgacgcaatc gttgaaggga 3060  
 aagttctcat tgggtccagt gacgcagccg aaggcgcaat tc 3102  
 45  
 <210> 18  
 <211> 3106  
 <212> DNA  
 <213> new AAV serotype, clone 7-2  
 50  
 <400> 18  
 gcggccgga attcgccctt cgagagacc aaagttcaac tgaacgaat cagccggttt 60  
 attgattaac aagtaattac aggttacggg tgaggtaacg ggtgccaatg gggcgaggct 120  
 55 cagtataaac ccttcgttg ttgacagca attccacatt attagacttg gcataatttg 180

EP 1 310 571 A2

	aggtgtactg aatctctgga ttccagcggt tgctgttttc tttctgcagt tcccactcga	240
5	tctccacgct gacctggccg gtgctgtact gcgtgataaa tgaggcaaac ttggcaggag	300
	taaacacctc tggaggatta gcaggtaccg ggggtgtttt gatgagaatt tgaggaggcg	360
	gggtgtttgag tccaaatccg cccatcaggg gagacgggtg aaagttgccg tccgtgtgag	420
10	gaatltttggc ccagatggga ccttgcaggt acacgtcccg gttctgccag accatgccgg	480
	gcagagcccc ctggctgttg acagtctgtg tctgggttcc ggccgtagac gattgcaggt	540
	tgctggagac cacaccgtat tcttctgtag ccacgggatt ggtggttttg atctcctcct	600
15	cgctggctcat tagcacgttt tccagcgttg tcttgttggc agccccggtt ttgccaaaaa	660
	ccagcactcc gttgatggga aagaactggg cctcgtcgtc cttgttgggt gccatggcta	720
	cggccgggtt ggttaatgaa tttctacat tcagatggta tttagtggcc ccgggtccagg	780
20	caaagttact gttgttgttg ctgtctatgt tttttgacag tctctgctgc cgataacagg	840
	gtccggggcag ccagttcttt gattgctcgg ccatgggtgtt gggcccagcc tgatggaact	900
	gcagctccct tgtggacccc gtagtgctct gggtcgggc caggtagtac aggtactggt	960
25	cgatgagggg attcatcagc cggtcacggc tctggctgtg cgcatagctg ctgtggaaag	1020
	gcacttcctc aaaggtgtag ctgaattcaa agttatcgcc cgttctcagc atctgagaag	1080
	gaaagtactc caggcagtag aaggaggaac gtcccacaga ctgactgccg ttgttttagag	1140
30	tcagatatcc gtactgagga atcatgaaca cgctccgagg gaacggaggg aggcagccct	1200
	gggtgcgcaga gccgaggacg tacggcagtt ggtactccga gtccgagaag acctgaatcg	1260
	tgctggtaag gttattagcg atggctgtaa cgccgtcgtt cgctgtgacc tccttgacct	1320
35	ggatgttgaa caacttgaac cgcagctttc tgggcccggaa tcccagttg ttgttgatga	1380
	gtcgttgcca gtcacgtggg gagaagtggc agtggaatct gttgaagtca aaatagcccc	1440
	aggggggtgct gtagccgaag aagtggttgt cgttggtagc cccgctctga cttgatatct	1500
40	gcttgtagag gtggttgttg taggtgggca gggcccagg gtgggtgctg gtgggtgatga	1560
	ctctgtcgcc cagccatgtg gaatcgcaat gccaatltcc ggaggcatta cccactccgt	1620
	cggcgccttc gttattgtct gccattgggt cgccaccgcc tgcagccatt gtaccagatc	1680
45	ccagacctga gggcgcgggc ggagggttctc cgagaggttg ggggtcgggc actgactctg	1740
	agtcgccagt ctgccc aaag ttgagcttct ttttagcggg cggctggccg ttcttgccga	1800
	tgcccgtgga ggagtcgggg gattctatgg gtctcttctt tccaggagcc gtcttagcgc	1860
50	cttctcaac cagaccgaga ggttcgagaa cccgcttctt ggcttgaag actgctcgcc	1920
	cgagggttgc cccaaaagac gtatcttctt gaagacgctc ctgaaactcg gcgtcggcgt	1980
	ggttgtactt gaggtacggg ttgtccccct gctcgagctg cttgtcgtag gccttgtcgt	2040
55	gctcgagggc cggggcgctc gcctcgttga ccggctctcc cttgtcgagt ccgttgaagg	2100

# EP 1 310 571 A2

	gtccgaggta cctgtagcca ggaagcacca gaccccggcc gtggtcctgc ttttgctggt	2160
5	tggtcttggg tttcggggct ccagggtttca agtcccacca ctgcgaatg ccctcagaga	2220
	ggttgccttc gagccaatct ggaagataac catcggcagc catacctggt ttaagtcatt	2280
	tattgctcag aaacacagtc atccagggtcc acgttggcca gatcgaggc cgagcaagca	2340
10	atctcgggag cccgccccag cagatgatga atggcacaga gtttccgata cgtcctcttt	2400
	ctgacgaccg gttgagattc tgacacgccg gggaaacatt ctgaacagtc tctggtcccg	2460
	tgcgtgaagc aaatgttgaa attctgattc attctctcgc atgtcttgca ggggaacagc	2520
15	atctgaagca tgcccgcggtg acgagaacat ttgttttggg acctgtcggc aaagtccacc	2580
	ggagctcctt ccgcgtctga cgtcgatgga tccgcgactg aggggcaggc ccgcttgggc	2640
	tcgcttttat ccgcgtcatc gggggcgggg ctcttgttgg ctccaccctt tctgacgtag	2700
20	aactcatacg ccacctcggg cacgtgatcc tgcgccagc ggaagaactc tttgacttcc	2760
	tgctttgtca ccttgccaaa gtcattgctc agacggcggg tgagttcaaa tttgaacatc	2820
	cggctcttga acggctgctg gtgctcgaag gtggtgctgt tcccgtaaat cacggcgcac	2880
25	atgttggtgt tggaagtga gatcacgggg gtgggatcga tctgggcgga cgacttgac	2940
	ttttggtcca cgcgcacctt gctgccgccg agaattggcct tggcggactc cacgacctg	3000
	gccgtcatcc tgccctctc ccaccagatc accatcttgt cgacgcaatc gttgaaggga	3060
30	aagttctcat tgggtccagtt gacgcagccg tagaaagggc gaattc	3106
	<210> 19	
	<211> 3105	
	<212> DNA	
35	<213> new AAV serotype, clone C1	
	<400> 19	
	gaattcgccc ttgctgcgtc aactggacca atgagaactt tcccttcaac gattgcgtcg	60
40	acaagatggt gatctggtgg gaggagggca agatgaccgc caaggctcgtg gagtccgcca	120
	aggccattct gggcggaagc aagggtgcgcg tggacaaaa gtgcaagtca tcggcccaga	180
	tcgacccac gcccgatgac gtcacctcca acaccaacat gtgcgccgtg atcgacggga	240
45	acagcaccac cttcgagcac cagcagccgc tgcaggaccg catgttcaag ttcgagctca	300
	cccgccgtct ggagcacgac tttggcaagg tgaccaagca ggaagtcaaa gagttcttcc	360
	gctgggtcga ggatcacgtg actgaggtgg cgcattgatt ctacgtcaga aaggcgagg	420
50	ccacaaaaag acccgcccc agtgacgcgg atataagcga gcccaagcgg gcctgcccct	480
	cagttgcgga gccatcgacg tcagacgcgg aagcaccggt ggactttgcg gacaggtacc	540
	aaaacaaatg ttctcgtcac gcgggcatgc ttcagatgct gtttccctgc aagacatgcg	600
55	agagaatgaa tcagaatttc aacgtctgct tcacgcacgg ggtcagagac tgctcagagt	660
	gcttccccgg cgcgtcagaa tctcaacctg tcgtcagaaa aaagacgtat cagaaactgt	720

# EP 1 310 571 A2

5	gcgcgattca	tcattctgctg	gggcggggcac	ccgagattgc	gtgttcggcc	cgcatctctg	780
	tcaacgtgga	cttgatgac	tgtgtttctg	agcaataaat	gacttaaacc	aggtatggct	840
	gctgacgggt	atcttccaga	ttggctcgag	gacaacctct	ctgagggcat	tcgcgagtgg	900
10	tgggacctga	aacctggagc	ccccaggccc	aaggccaacc	agcagaagca	ggacgacggc	960
	cggggtctgg	tgcttcctgg	ctacaagtac	ctcggaccct	tcaacggact	cgacaagggg	1020
	gagcccgtca	acgcggcgga	cgcagcggcc	ctcagacacg	acaaggccta	cgaccagcag	1080
15	ctcaaagcgg	gtgacaatcc	gtacctgcgg	tataaccacg	ccgacgccga	gtttcaggag	1140
	cgtctgcaag	aagatacgtc	ttttgggggc	aacctcgggc	gagcagtctt	ccaggccaag	1200
	aagagggtag	tcgaacctct	gggcctgggt	gaagaagggt	ctaagacggc	tcctggaaag	1260
20	aagagaccgt	tagagtcacc	acaagagccc	gactcctcct	caggaatcgg	caaaaaaggc	1320
	aaacaaccag	ccaaaaagag	actcaacttt	gaagaggaca	ctggagccgg	agacggaccc	1380
	cctgaaggat	cagataccag	cgccatgtct	tcagacattg	aatgctgtgc	agcaccgggc	1440
25	ggaaatgctg	tcgatgcggg	acaaggttcc	gatggagtgg	gtaatgcctc	gggtgattgg	1500
	cattgctgatt	ccacctggtc	tgagggcaag	gtcacaacaa	cctcgaccag	aacctgggtc	1560
	ttgcccacct	acaacaacca	cttgtacctg	cggctcggaa	caacatcaaa	cagcaacacc	1620
30	tacaacggat	tctccacccc	ctggggatac	tttgacttta	acagattcca	ctgtcacttc	1680
	tcaccacgtg	actggcaaa	actcatcaac	aacaactggg	gactacgacc	aaaagccatg	1740
	cgcggttaaaa	tcttcaatat	ccaagttaag	gaggtcacaa	cgtcgaacgg	cgagactacg	1800
35	gtcgctaata	accttaccag	cacggttcag	atatttgctg	actcgtcgta	tgagctcccg	1860
	tacgtgatgg	acgctggaca	agagggaagt	ctgtctcctt	tcccgaatga	cgtcttcatg	1920
	gtgcctcaat	atggctactg	tggcattgtg	actggcgaaa	atcagaacca	gacggacaga	1980
40	aatgctttct	actgcctgga	gtattttcct	tcacaaatgc	tgagaactgg	caataacttt	2040
	gaaatggctt	acaactttgg	gaaggtgccg	ttccactcaa	tgtatgctta	cagccagagc	2100
	ccggacagac	tgatgaatcc	cctcctggac	cagtacctgt	ggcacttaca	gtcgaccacc	2160
45	tctggagaga	ctctgaatca	aggcaatgca	gcaaccacat	ttggaaaaat	caggagtggga	2220
	gactttgcct	tttacagaaa	gaactggctg	cctgggcctt	gtgttaaaca	gcagagactc	2280
	tcaaaaactg	ccagtcaaaa	ttacaagatt	cctgccagcg	ggggcaacgc	tctgttaaag	2340
50	tatgacaccc	actatacctt	aaacaaccgc	tggagcaaca	tagcgctcgg	acctccaatg	2400
	gcaacagctg	gaccttcaga	tggggacttc	agcaacgccc	agctcatctt	ccctggacca	2460
	tcagtcaccg	gaaacacaac	aacctcagca	aacaatctgt	tgttttacatc	agaagaagaa	2520
55	attgctgcca	ccaacccaag	agacacggac	atgtttggtc	agattgctga	caataatcag	2580
	aatgctacaa	ctgctcccat	aaccggcaac	gtgactgcta	tgggagtgtc	tcctggcatg	2640

# EP 1 310 571 A2

5           gtgtggcaaa acagagacat ttactaccaa gggccaattt gggccaagat cccacacgcg   2700

          gacggacatt ttcataccttc accgctaatt ggcggttttg gactgaaaca tccgcctccc   2760

          cagatatatta tcaaaaacac ccccgtaacct gccaatcctg cgacaacctt cactgcagcc   2820

10          agagtggact ctttcatcac acaatacagc accggccagg tcgctgttca gattgaatgg   2880

          gaaatcgaaa aggaacgctc caaacgctgg aatcctgaag tgcagtttac ttcaaactat   2940

          gggaaccagt cttctatggt gtgggctccc gatacaactg ggaagtatac agagccgcgg   3000

15          gttattggct ctcgttattt gactaatcat ttgtaactgc ctagttaatc aataaaccgt   3060

          gtgattcggt tcagttgaac tttggtctct gcgaaggcg aattc                   3105

20          <210> 20  
          <211> 3105  
          <212> DNA  
          <213> new AAV serotype, clone C3

          <400> 20  
          gaattcgccc ttgctgcgtc aactggacca atgagaactt tcccttcaac gattgcgtcg   60

25          acaagatggg gatctgggtg gaggagggca agatgaccgc caaggctgtg gagtccgccca   120

          aggccattct gggcggaagc aagggtgcgc tggacaaaaa gtgcaagtca tcggcccaga   180

          tcgacccac gcccgatgc gtcacctcca acaccaacat gtgcgccgtg atcgacggga   240

30          acagcaccac cttcgagcac cagcagccgc tgcaggaccg catgttcaag ttcgagctca   300

          cccgcctct ggagcacgac tttggcaagg tgaccaagca ggaagtcaaa gagttcttcc   360

          gctgggtca ggatcacgtg actgaggtg cgcattgagt ctacgtcaga aaggggcgag   420

35          ccacaaaaag acccgcccc agtgacgcgg atataagcga gcccaagcgg gcctgcccct   480

          cagttgcgga gccatcgacg tcagacgcgg aagcaccggt ggactttgcg gacagggtacc   540

          aaaacaaatg ttctcgtcac gcgggcatgc ttcagatgct gtttccttgc aagacatgcg   600

40          agagaatgaa tcagaatttc aacgtctgct tcacgcacgg ggtcagagac tgctcagagt   660

          gcttccccgg cgcgtcagaa tctcaacctg tcgtcagaaa aaagacgtat cagaaactgt   720

          gcgcgattca tcatctgctg gggcgggcac ccgagattgc gtgttcggcc tcgcatctcg   780

45          tcaacgtgga cttggatgac tgtgtttctg agcaataaat gacttaaacc aggtatggct   840

          gctgacggtt atcttccaga ttggctcgag gacaacctct ctgagggcac tcgcgagtgg   900

          tgggacctga aacctggagc cccaagctc aaggccaacc agcagaagca ggacgacggc   960

50          cgggggtctg tgcttcttgg ctacaagtac ctcgaccct tccacggact cgacaagggg   1020

          gagcccgta acgcggcgga cgcagcgccc ctcgagcacg acaaggccta cgaccagcag   1080

          ctcaaagcgg gtgacaatcc gtacctgcgg tataaccacg ccgacgccga gtttcaggag   1140

55          cgtctgcaag aagatacgtc ttttgggggc aacctcgggc gagcagtcct ccaggccaag   1200

EP 1 310 571 A2

	aagagggtac tcgaaccact gggcctgggt gaagaagggtg ctaagacggc tcctggaaag	1260
	aagagaccgt tagagtcacc acaagagccc gactcctcct caggaatcgg caaaaaaggc	1320
5	aaacaaccag ccaaaaagag actcaacttt gaagaggaca ctggagccgg agacggaccc	1380
	cctgaaggat cagataccag cgccatgtct tcagacattg aaatgcgtgc agcaccgggc	1440
	ggaaatgctg tcgatgcggg acaagggttc gatggagtgg gtaatgcctc ggggtgattg	1500
10	cattgcgatt ccacctggtc tgagggaag gtcacaacaa cctcgaccag aacctgggtc	1560
	ttgcccacct acaacaacca cttgtacctg cggctcggaa caacatcaaa cagcaacacc	1620
	tacaacggat tctccacccc ctggggatac tttgacttta acagattcca ctgtcacttc	1680
15	tcaccacgtg actggcaaa actcatcaac aacaactggg gactacgacc aaaagccatg	1740
	cgcgtaaata tcttcaatat ccaagttaag gaggtcacia cgtcgaacgg cgagactacg	1800
	gtcgctaata accttaccag cacgggttcag atatttgagg actcgtcgta tgagctcccg	1860
20	tacgtgatgg acgctggaca agagggaagt ctgcctcctt tcccgaatga cgtcttcattg	1920
	gtgcctcaat atggctactg tggcattgtg actggcgaaa atcagaacca gacggacaga	1980
	aatgctttct actgcctgga gtattttcct tcacaaatgc tgagaactgg caataacttt	2040
25	gaaatggctt acaactttga gaagggtgcc ttccactcaa tgtatgctca cagccagagc	2100
	ctggacagac tgatgaatcc cctcctggac cagtacctgt ggcaattaca gtcgaccacc	2160
	tctggagaga ctctgaatca aggcaatgca gcaaccacat ttggaaaaat caggagtggga	2220
30	gactttgcct ttacagaaa gaactggctg cctgggcctt gtgttaaaca gcagagattc	2280
	tcaaaaactg ccagtcaaaa ttacaagatt cctgccagcg ggggcaacgc tctgttaaag	2340
	tatgacaccc actatacctt aaacaaccgc tggagcaaca tagcgctgg acctccaatg	2400
35	gcaacagctg gaccttcaga tggggacttc agcaacgccc agctcatctt ccctggacca	2460
	tcagtcaccg gaaacacaac aacctcagca aacaatctgt tgtttacatc agaaggagaa	2520
	attgctgcca ccaacccaag agacacggac atgtttggtc agattgctga caataatcag	2580
40	aatgctacaa ctgctcccat aaccggcaac gtgactgcta tgggagtgtt tcctggcatg	2640
	gtgtggcaaa acagagacat ttactaccaa gggccaattt gggccaagat cccacacggc	2700
	gacggacatt ttcctccttc accgctaatt ggcggttttg gactgaaaca tccgcctccc	2760
45	cagatattta tcaaaaacac ccccgctacct gccaatcctg cgacaacctt cactgcagcc	2820
	agagtggact ctttcatcac acaatacagc accggccagg tcgctgttca gattgaatgg	2880
	gaaatcgaaa aggaacgctc caaacgccgg aatcctgaag tgcagtttac ttcaaactat	2940
50	gggaaccagt cttctatgtt gtgggctccc gatacaactg ggaagtatac agagccgagg	3000
	gttattggct ctcgttattt gactaatcat ttgtaactgc ctagttaatc aataaaccgt	3060
55	gtgattcggt tcagttgaac tttgggtctt gcgaagggag aattc	3105

# EP 1 310 571 A2

<210> 21  
 <211> 3105  
 <212> DNA  
 5 <213> new AAV serotype, clone C5  
  
 <400> 21  
 gaattcgccc ttgcgagaga ccaaagttca actgaaacga atcacacggt ttattgatta 60  
 actaggcagt tacaaatgat tagtcaaata acgagagcca ataaccgcg gctctgtata 120  
 10 cttcccagtt gtatcgggag cccacaacat agaagactgg ttcccacagt ttgaagtaaa 180  
 ctgcacttca ggattccagc gtttgaggcg ttccttttcg atttcccatt caatctgaac 240  
 agcgacctgg ccggtgctgt attgtgtgat gaaagagtcc actctggctg cagtgaagggt 300  
 15 tgtcgcagga taggcaggtg cgggggtgtt ttgataaat atctggggag gcggatgttt 360  
 cagtccaaaa ccgccaatta gcggtgaagg atgaaaatgt ccgtccgcgt gtgggatctt 420  
 ggcccaaatt ggcccttggg agtaaagtgc tctgttttgc cacaccatgc caggaagcac 480  
 20 tcccatagca gtcacgttgc cggttatggg agcagttgta gcattctgat tattgtcagc 540  
 aatctgacca aacatgtccg tgtctcttgg gttggtggca gcaatttctt cttctgatgt 600  
 25 aaacaacaga ttgtttgctg aggttggtgt gtttccgggtg actgatggtc caggggaagat 660  
 gagctgggag ttgctgaagt ccccatctga aggtccagct gttgccattg gaggtccagg 720  
 cgctatgttg ctccagcggg tgtttaaggt atagtgggtg tcatacttta acagagcgtt 780  
 30 gccccgcgtg gcaggaatct tgtaattttg actggcagtt tttgagaatc tctgctgttt 840  
 aacacaaggc ccaggcagcc agttctttct gtaaaaggca aagtctccac tcctgatttt 900  
 tccaaatgtg gttgctgcat tgccttgatt cagagtctct ccagaggtgg tcgactgtaa 960  
 35 gtgccacagg tactggtcca ggaggggatt catcagtcctg tccaggctct ggctgtgagc 1020  
 atacattgag tggaacggca ccttctcaaa gttgtaagcc gtttcaaagt tattgccagt 1080  
 tctcagcatt tgtgaaggaa aatactccag gcagtagaaa gcatttctgt ccgtctggtt 1140  
 40 ctgattttcg ccagtcacaa tgccacagta gccatattga ggcaccatga agacgtcatt 1200  
 ggggaaagga ggcagacttc cctcttgtcc agcgtccatc acgtacggga gtcatacga 1260  
 cgagtccgca aatatctgaa ccgtgctggt aagggtatta gcgaccgtag tctcgccgtt 1320  
 45 cgacgttgtg acctccttaa cttggatatt gaagatttta acgcgcattg cttttggtcg 1380  
 tagtccccag ttgttgttga tgagtctttg ccagtcacgt ggtgagaagt gacagtggaa 1440  
 tctgttaaag tcaaagtatc ccaggggggt ggagaatccg ttgtaggtgt tgctgtttga 1500  
 50 tgttgttccg agccgcaggt acaagtgggt gttgtaggtg ggcaagacc aggttctggt 1560  
 cgaggttggt gtgaccttgc cctcagacca ggtggaatcg caatgccaat cacccgaggc 1620  
 attaccact ccacggaac cttgtccgc atcgacagca tttccgccc gtgctgcacg 1680  
 55 catttcaatg tctgaagaca tggcgctggt atctgatcct tcaggggggtc cgtctccggc 1740

# EP 1 310 571 A2

	tccagtgtcc tcttcaaagt tgagtctctt tttggctggt tgtttgcctt ttttgccgat	1800
5	tcctgaggag gaggcgggct cttgtggtga ctctaacggg ctcttctttc caggagccgt	1850
	cttagcacct tcttcaacca ggcccagagg ttcgagtacc ctcttcttgg cctggaagac	1920
	tgctcgcccc aggttgcccc caaaagacgt atcttcttgc agacgctcct gaaactcggc	1980
10	gtcggcggtg ttataccgca ggtacggatt gtcacccgct ttgagctgct ggtcgtaggc	2040
	cttgctcgtc tcgagggccg ctgcgtccgc cgcgttgacg ggctccccct tgtcgagtcc	2100
	gttgaagggt ccgaggtact cgtagccagg aagcaccaga ccccgccgt cgtcctgctt	2160
15	ctgctggttg gccttgggct tgggggctcc aggtttcagg tcccaccact cgcaatgcc	2220
	ctcagagagg ttgtcctcga gccaatcttg aagataaccg tcagcagcca tacctggttt	2280
	aagtcattha ttgctcagaa acacagtcac ccaagtccac gttgacgaga tcgcaggccg	2340
20	aacacgcaat ctcggttgcc cggcccagca gatgatgaat cgcgcacagt ttctgatacg	2400
	tcttttttct gacgacgggt tgagattctg acgcgcggg gaagcactct gagcagtctc	2460
	tgaccccggt cgtgaagcag acgttgaaat tctgattcat tctctcgcat gtcttgacgg	2520
25	gaaacagcat ctgaagcatg ccgcggtgac gagaacattt gttttggtac ctgtccgcaa	2580
	ggtccaccgg tgcttcgcgc tctgacgtcg atggctccgc aactgagggg caggcccgt	2640
	tgggctcgct tatatccgcg tcaactgggg cgggtctttt ggtggctccg ccctttctga	2700
30	cgtagaactc atgcgccacc tcagtcacgt gatectgagc ccagcggaag aactctttga	2760
	cttctgctt ggtcaccttg ccaaagtcgt gctccagacg gcgggtgagc tcgaacttga	2820
	acatgcggtc ctgcagcggc tgctgggtgct cgaagggtgg gctgttcccg tcgatcacgg	2880
35	cgcacatggt ggtgttgag gtgacgatca cgggcgtggg gtcgatctgg gccgatgact	2940
	tgcacttttg gtccacgcgc accttgcttc cggccagaat ggccttggcg gactccacga	3000
	ccttggcggg catcttgccc tctcccacc agatcaccat cttgtcgacg caatcgttga	3060
40	agggaaagtt ctatttggtc cagttgacgc agcaaggcg aattc	3105
	<210> 22	
	<211> 3094	
45	<212> DNA	
	<213> new AAV serotype, clone F1	
	<400> 22	
	gaattcggcc ttgctgcgtc aactggacca agagaacttt cccttcaacg attgcgtcga	60
50	caagatggtg atctggtggg aggagggcaa gatgacggcc aaggctcgtg agtccgcaa	120
	agccattctg ggcggaagca aggtgcgcgt cgacaaaaag tgcaagtcct cggcccagat	180
	cgatcccacc cccgtgatcg tcacctcaa caccaacatg tgcgccgtga tcgacgggaa	240
55	cagcaccacc ttcgagcacc agcagccgtt gcaggaccgg atgttcaaat ttgaactcac	300

EP 1 310 571 A2

	ccgccgtctg gaacacgact ttggcaaggt gaccaagcag gaagtcaaag agttcttccg	360
	ctgggctagt gatcacgtga ctgaggtgac gcatgagttc tacgtcagaa agggcggagc	420
5	cagcaaaaga cccgcccccg atgacgcgga tataagcgag cccaagcggg cctgtccctc	480
	agtcacggac ccacgcagct cagacgcgga aggagctccg gtggactttg ccgacaggta	540
	ccaaaacaaa tgttctcgtc acgcgggcat gcttcagatg ctgtttccct gcaaacgtg	600
10	cgagagaatg aatcagaatt tcaacatttg cttcacgcac ggggtcagag actgtttaga	660
	atgtttcccc ggcgtgtcag aatctcaacc ggtcgtcaga aaaaagacgt atcggaagct	720
	gtgtgcgatt catcatctgc tggggcgggc acccgagatt gcttgctcgg cctgcgacct	780
15	ggtcaacgtg gacctggacg actgtgtttc tgagcaataa atgacttaaa ccgggtatgg	840
	ctgccgatgg ttatcttcca gattggctcg aggacaacct ctctgagggc attcgcgagt	900
	ggtgggacct gaaacctgga gccccgaac ccaaagccaa ccagcaaaag caggacgacg	960
20	gccggggctc ggtgcttcct ggctacaagt acctcggacc cttcaacgga ctgcacaagg	1020
	gggagcccg tcaacgcggc gacgcagcgg cctcgcagca cgacaaggcc tacgaccagc	1080
	agctcaaagc gggtgacaat ccgtacctgc ggtataacca cgccgacgcc gagtttcagg	1140
25	agcgtctgca agaagatacg tcatttgggg gcaacctcgg gcgagcagtc ttccaggcca	1200
	agaagcgggt tctcgaacct ctcggtctgg ttgaggaagg cgctaagacg gctcctgga	1260
	agaagagacc catagactct ccagactcct ccacgggcat cggcaaaaaa ggccagcagc	1320
30	ccgctaataa gaagctcaat tttggtcaga ctggcgactc agagtcagtc cccgaccctc	1380
	aacctcttgg agaacctcca gcagcgccct ctagtgtggg atctggtaca atggctgcag	1440
	gcgggtgggc accaatggca gacaataacg aagggtgccg cggagtgggt aatgcctcag	1500
35	gaaattggca ttgcgattcc acatggctgg gcgacagagt catcaccacc agcaccagaa	1560
	cctgggccct cccacctac aacaaccacc tctacaagca aatctccagc agcagctcag	1620
	gagccaccaa tgacaaccac tacttcggct acagcaccct ctgggggtat tttgacttta	1680
40	acagattcca ctgccacttc tcaccacgtg actggcagcg actcatcaac aacaactggg	1740
	gattccggcc caagaagctg cggttcaagc tcttcaacat ccaggtaag gaggtcacia	1800
	cgaatgacgg cgtcacgacc atcgctaata accttaccag cacgggtcag gtcttctcgg	1860
45	actcggaata ccagctgccg tacgtcctcg gctctgcgca ccagggtcgc ctgcctcgt	1920
	tcccggcgga cgtcttcatg attcctcagt acggctacct gactctgaac aacggcagcc	1980
	aatcgggtggg ccgttctctc ttctactgcc tggaatatct cccctctcaa atgctgagaa	2040
50	cgggcaacaa ctttgagttc agttacagct tcgaggacgt gcctttccac agcagctacg	2100
	cgcacagcca gagcctagac cggctgatga accctctcat cgaccagtac ctgtactacc	2160
55	tggccccggac ccagagcacc acgggttcca ccagggaact gcaatttcat caagctgggc	2220

# EP 1 310 571 A2

	ccaatactat ggccgagcag tcaaagaact ggctgcctgg accctgctat aggcaacagg	2280
	gactgtcaaa gaacttggac ttaacaaca acagcaattht tgcttggact gctgccacta	2340
5	aatatcatct gaatggcaga aactctttga ccaatcctgg cattcccatg gcaaccaaca	2400
	aggatgatga ggaccagttc tttcccatca acgggggtact ggthttttggc aagacgggag	2460
	ctgccaacaa aactacgctg gaaaacgttc tgatgaccag cgaggaggag atcaagacca	2520
10	ctaaccctgt ggctacagaa gaatacgggtg tggctctccag caacctgcag ccgtctacag	2580
	ccgggcctca atcacagact atcaacagcc agggagcact gcctggcatg gtctggcaga	2640
	accgggacgt gtatctgcag ggtcccatct gggccaaaat tcctcacacg gatggcaact	2700
15	ttcacccgtc tcctctgatg ggcgggttttg gactcaaaaa cccgcctcca cagatcctga	2760
	tcaaaaacac acctgtacct gctaatactc cggagggtgtt tactcctgcc aagtttgcct	2820
	ccttcatcac gcagtacagc accggacaag tcagcgtgga aatcgagtgg gagctgcaga	2880
20	aagaaaacag caagcgtctg aaccagaaaa ttcagtatac ttccaattat gccaaagtcta	2940
	ataatgttga atttgctgtg aacctgtatg gtgtttatac tgagcctcgc cccattggca	3000
	ctcgttacct ccccgtaat ctgtaattgc ttgttaatca ataaaccggg tgattcgttt	3060
25	cagttgaact ttggtctctg cgaaggcgca attc	3094
30	<210> 23	
	<211> 3095	
	<212> DNA	
	<213> new AAV serotype, clone F3	
35	<400> 23	
	gaattcgccc ttgcgagaga ccaaagttca actgaaacga atcaaccggg ttattgatta	60
	acaagcaatt acagattacg ggtgaggtaa cgagtgccaa tggggcgagg ctcaagtataa	120
	acaccatcag ggttcacagc aaattcaaca ttattagact tggcataatt ggaagtatac	180
40	tgaattttctg gggtccagcg cttgctgttt tctttctgca gctcccactc gatttccacg	240
	ctgacttgtc cgggtgctgta ctgcgtgatg aaggaggcaa acttggcagg agtaaacacc	300
	tccggaggat tagcaggtag aggtgtgttt ttgatcagga tctgtggagg cgggtgtttg	360
45	agtcctaaaac cgcccatcag aggagacggg tgaaagttgc catccgtgtg aggaattttg	420
	gccagatgg gacctgcag atacacgtcc cggttctgcc agaccatgcc aggcagtgtc	480
	ccctggctgt tgatagtctg tgattgaggc ccggctgtag acgactgcag gttgctggag	540
50	accacaccgt attcttctgt agccacaggg ttagtgggtc tgatctctc ctcgctggtc	600
	atcagaacgt tttccagcgt agttttgttg gcagctcccg tcttgccaaa aaccagtacc	660
	ccgttgatgg gaaagaactg gtcctcatca tcctgtttgg ttgccatggg aatgccagga	720
55	ttggtcaaag agtttctgcc attcagatga tatttagtgg cagcagttca ggcaaaattg	780

EP 1 310 571 A2

	ctgttggtgt taaagtccaa gttctttgac agtctctgtt gcctatagca ggggtccaggc	840
	agccagttct ttgactgctc ggccatagta ttggggccag cttgatgaaa ttgcagttcc	900
5	ctggtggaac ccgtggtgct ctgggtccgg gccaggtagt acaggtagtg gtcgatgaga	960
	gggttcatca gccgggtctag gctctggctg tgcgcgtagc tgctgtggaaggcacgtcc	1020
	tcgaagctgt aactgaactc aaagtgtgtt cccgttctca gcatttgaga ggggaaatat	1080
10	tccaggcagt agaaggagga acggccacc gattggctgc cgttggtccag agtcaggtag	1140
	ccgtactgag gaatcatgaa gacgtccgcc gggaacggag gcaggcagcc ctggtgcgca	1200
	gagccgagga cgtacggcag ctggtattcc gagtccgaga agacctgaac cgtgctggta	1260
15	aggttattag cgatgggtcgt gacgccgtca ttcgttggtga cctccttgac ctggatgttg	1320
	aggagcttga accgcagctt ctggggccgg aatccccagt tgttgttgat gagtgcgtgc	1380
	cagtcacgtg gtgagaagtg gcagtggaaat ctgttaaagt caaaataccc ccagggggtg	1440
20	ctgtagccga agtagtggtt gtcattgggtg gctcctgagc tgctgctgga gatttgcttg	1500
	tagagggtgtg tgttgtaggt ggggagggcc cagggttctgg tgctggtggt gatgactctg	1560
	tcgcccagcc atgtggaatc gcaatgccaa ttcctgagg cattaccac tccgtcggca	1620
25	ccttcgttat tgtctgccat tgggtgcgca ccgcctgcag ccattgtacc agatcccaca	1680
	ctagagggcg ctgctggagg ttctccaaga ggttgagggt cggggactga ctctgagtcg	1740
	ccagtctgac caaaattgag cttcttttta gcgggctgct ggctttttt gccgatgcc	1800
30	gtggaggagt ctggagagcc tatgggtctc ttctttccag gagccgtctt agcgccttcc	1860
	tcaaccagac cgagagggtc gagaaccgc ttcttggcct ggaagactgc tcgcccagg	1920
	ttgccccaa atgacgtatc ttcttgca gaactcctgaa actcggcgtc ggcgtggtta	1980
35	taccgcaggt acggattgtc acccgctttg agctgctggt cgtaggcctt gtcgtgctcg	2040
	agggccgctg cgtccgccgc gttgacgggc tccccctgt cgagtcggt gaagggtccg	2100
	aggtagcttg agccaggaag caccagaccc cggccgtcgt cctgcttttg ctggttggt	2160
40	ttgggtttcg gggctccagg ttccaggctc caccactcgc gaatgccctc agagaggttg	2220
	tcctcgagcc aatctggaag ataaccatcg gcagccatac ctggtttaag tcattttattg	2280
	ctcagaaaca cagtcgtcca ggtccacgtt gaccaggctg caggccgagc aagcaatctc	2340
45	gggtgcccgc ccagcagat gatgaatcgc acacagcttc cgatacgtct tttttctgac	2400
	gaccggttga gattctgaca cgccggggaa acattctaaa cagtctctga ccccgctgct	2460
	gaagcaaattg ttgaaattct gattcattct ctcgcacgtt ttgcaggga acagcacctg	2520
50	aagcatgccc gcgtgacgag aacatttggt ttggtacctg tcggcaagt ccaccggagc	2580
	tccttcgcgc tctgacgtcg atgggtccgt gactgaggga cgggccgct tgggtcgcct	2640
55	tatatccgcg tcatcggggg cgggtctttt gctggctccg ccctttctga cgtagaactc	2700

# EP 1 310 571 A2

	atgcgtcacc tcagtcacgt gatcactagc ccagcgggaag aactctttga cttcctgctt	2760
	tgtcaccttg ccaaagtcgt gttccagacg gcgggtgagt tcaaatttga acatccggtc	2820
5	ctgcaacggt tgctgggtgct cgaagggtggt gctgttcccg tcgatcacgg cgcacatggt	2880
	ggtgttggag gtgacgatca cgggggtggg atcgatctgg gcggacgact tgcacttttg	2940
	gtccacgcgc accttgctgc cgccgagaat ggccttggcg gactccacga ccttggccgt	3000
10	catcttgccc tcctcccacc agatcaccat cttgtcgacg caatcgttga agggaaagtt	3060
	ctcattggtc cagttgacgc agcaaggggc aattc	3095
15	<210> 24 <211> 3095 <212> DNA <213> new AAV serotype, clone F5	
20	<400> 24 gaattcgccc ttcgcagaga ccaaagttca actgaaacga atcaaccggt ttattgatta	60
	acaagcaatt acagattacg ggtgaggtaa cgagtgccaa tggggcgagg ctcagtataa	120
	acaccatcag ggttcacagc aaattcaaca ttattagact tggcataatt ggaagtatac	180
25	tgaatttctg ggttccagcg cttgctgttt tctttctgca gctcccactc gatttccacg	240
	ctgacttgtc cgggtctgta ctgcgtgatg aaggaggcaa acttggcagg agtaaacacc	300
	tccggaggat tagcaggtag aggtgtgttt ttgatcagga tctgtggagg cgggtgttcg	360
30	agtccaaaac cgcccatcag aggagacggg tgaaagttgc catccgtgtg aggaattttg	420
	gccagatgg gacctgcag atacacgtcc cggttctgcc agaccatgcc aggcagtgtc	480
	ccctggctgt tgatagtctg tgattgaggc cgggctgtag acgactgcag gttgctggag	540
35	accacaccgt attcttctgt agccacaggg ttagtggctc tgatctctc ctcgctggtc	600
	atcagaacgt tttccagcgt agttttgttg gcagctcccg tcttgccaaa aaccagtacc	660
	ccgttgatgg gaaagaactg gtcctcatca tccttgttgg ttgccatggg aatgccagga	720
40	ttggtcaaag agtttctgcc attcagatga tatttagtgg cagcagtcca ggcaaaattg	780
	ctgttggtgt taaagtccaa gttctttgac agtctctgtt gcctatagca ggggtccaggc	840
	agccagttct ttgactgctc ggccatagta ttgggccag cttgatgaaa ttgcagttcc	900
45	ctggtggaac ccgtgggtgct ctgggtccgg gccaggtagt acagggtactg gtcgatgaga	960
	gggttcatca gccggcttag gctctggctg tgcgcgtagc tgctgtggaa aggcacgtcc	1020
	tcgaagctgt aactgaactc aaagtgttg cccgttctca gcatttgaga ggggaaatat	1080
50	tccaggcagt agaaggagga acggcccacc gattggctgc cgttggtcag agtcaggtag	1140
	ccgtactgag gaatcatgaa gacgtccgcc gggaacggag gcaggcagcc ctggtgcgca	1200
	gagccgagga cgtacggcag ctggtattcc gagtccgaga agacctgaac cgtgctggta	1260
55	aggttattag cgatggtcgt gacgccgtca ttcgttgtga cctccttgac ctggatgttg	1320

# EP 1 310 571 A2

aagagcttga accgcagctt cttgggccgg aatccccagt tgttggtgat gagtcgctgc 1380  
 5 cagtcacgtg gtgagaagtg gcagtggaat ctgttaaagt caaaataccc ccaggggggtg 1440  
 ctgtagccga agtagtggtt gtcattgggtg gctcctgagc tgctgctgga gatttgcttg 1500  
 tagaggtggt tgttgtaggt ggggagggcc caggttctgg tgctggtggt gatgactctg 1560  
 10 tcgcccagcc atgtggaatc gcaatgccaa tttcctgagg cattaccac tccgtcggca 1620  
 ccttcgttat tgtctgccgt tgggtgcgca ccgctgcag ccattgtacc agatcccaca 1680  
 ctagagggcg ctgctggagg ttctccaaga ggttgagggt cggggactga ctctgagtcg 1740  
 15 ccagtctgac caaaattgag cttcttttta gcgggtgct gccctttttt gccgatgcc 1800  
 gtggaggagt ctggagagtc tatgggtctc ttctttccag gagccgtctt agcgcttcc 1860  
 tcaaccagac cgagaggttc gagaaccgc ttcttggcct ggaagactgc tcgcccagg 1920  
 20 ttgccccaa atgacgtatc ttcttgcagg cgctcctgaa actcggcgtc ggcgtggtta 1980  
 taccgcaggt acggattgtc acccgctttg agctgctggt cgtaggcctt gtcgtgctcg 2040  
 agggccgctg cgtccgcccgc gttgacgggc tccccctgt cgagtcctt gaagggtccg 2100  
 25 aggtacttgt agccaggaag caccagacc cgccgctcgt cctgcttttg ctggttggt 2160  
 ttgggtttcg gggtccagg ttctcaggtc caccactcgc gaatgcctc agagaggttg 2220  
 tcctcgagcc aatctggaag ataaccatcg gcagccatac ctggtttaag ccatttattg 2280  
 30 ctcaaaaca cagtcgtcca ggtccacgtt gaccaggtcg caggccgagc aggcaatctc 2340  
 ggggtgccgc ccagcagat gatgaatcgc acacagcttc cgatacgtct tttttctgac 2400  
 gaccggttga gattctgaca cgccggggaa acattctaaa cagtctctga ccccgctcgt 2460  
 35 gaagcaaattg ttgaaattct gattcattct ctgcacgtt ttgcaggga acagcatctg 2520  
 aagcatgccc gcgtggcgag aacatttggt ttggtacctg tcggcaaagt ccaccggagc 2580  
 tccttcgcg tetgacgtcg atgggtccgt gactgaggga caggcccgct tgggtcgtc 2640  
 40 tatatccgcg tcatcgggg cggttctttt gctggctccg ccctttctga cgtagaactc 2700  
 atgcgtcacc tcagtcacgt gatcactagc ccagcggaag aactctttga cttcctgctt 2760  
 tgtcaccttg ccaaagtcgt gttccagacg gcgggtgagt tcaaatttga acatccggtc 2820  
 45 ctgcaacggc tgctggtgct cgaaggtggt gctgttcccg tcgatcacgg cgcgcattgt 2880  
 ggtgttgag gtgacgatca cgggggtggg atcgatctgg gcggacgact tgcacttttg 2940  
 gtccacgccc accttgctgc cgccgagaat ggcttggcg gactccacga ccttgccgt 3000  
 50 catcttgccc tcctccacc agatcaccat cttgtcgacg caatcgttga agggaaagt 3060  
 ctcatgtgct cagttgacgc agcaaggcg aattc 3095

55

# EP 1 310 571 A2

<210> 25  
 <211> 3142  
 <212> DNA  
 <213> new AAV serotype, clone H6

5

<400> 25  
 aaaacgacgg gccagtgatt gtaatacgac tcactatagg gcgaaattga aattagcggc 60  
 cgccaattcg cctttcgcag agaccaaagt tcaactgaaa cgaattaaac ggttttattga 120  
 10 ttaacaagca attacagatt acgagtcagg tatctggtgc caatggggcg aggctctgaa 180  
 tacacacccat tagtgtccac agtaaagtcc acattaacag acttggtgta gttggaagtg 240  
 tactgaattt cgggattcca gcgtttgctg ttctccttct gcagctccca ctcgatctcc 300  
 15 acgctgacct gtcccgtgga atactgtgtg atgaaagaag caaacttggc agaactgaag 360  
 tttgtgggag gattggctgg aacgggagtg tttttgatca tgatctgagg aggcgggtgt 420  
 ttgagtccaa aacctcccat cagtggagaa ggatgaaagt gtccatcggg gtgaggaatc 480  
 20 ttggcccaaa tgggtccctg caggtacacg tctcgatcct gccacacccat accaggtaac 540  
 gtccttgggt gattgacagt tccagtagtt ggaccagtgt ttgagttttg caaattatct 600  
 gacacagtcc cgtactgctc cgtagccacg ggattgggtg ccctgatttc ttcttcatct 660  
 25 gtaatcatga cattttccaa atccgcgtcg ttggcatttg ttcttgttt accaaatctc 720  
 agggttccat gcatggggaa aaacttttct tcgtcatcct tgtgactggc catagctggt 780  
 cctggattaa ccaacgagtc ccggccattt agatgatact ttgtagctgc agtccaggga 840  
 30 aagttgctgt tgttgtgtgc gtttgctgtt tttgacagac gctgctgtct gtagcaagggt 900  
 ccaggcagcc agtttttagc ttgaagagac atgttggttg gtccagcttg gctaaacagt 960  
 agccgagact gctgaagagt tccactatct gtttgtgtct tgttcagata atacaggtac 1020  
 35 tggtcgatca gaggattcat cagccgatcc agactctggc tgtgagcgta gctgctgtgg 1080  
 aaaggcacgt cttcaaaagt gtagctgaac tgaaagttgt ttccagtacg cagcatctga 1140  
 40 gaaggaaagt actccaggca gtaaaaggaa gagcgtccta ccgcctgact cccgttggtc 1200  
 agggtgaggt atccatactg tgggaccatg aagacgtccg ctggaaacgg cgggagccat 1260  
 ccttgatgcy ccgagcccag gacgtacggg agctgggtact ccgagtcagt aaacacctga 1320  
 45 accgtgctgg taaggttatt ggcaatcgtc gtcgtaccgt cattctgctg gacctctttg 1380  
 acttgaatat taaagagctt gaagttgagt cttttgggcc ggaatccccg gttgttgttg 1440  
 acgagtcttt gccagtcacg tgggtgaaaag tggcagtggg atctgttgaa gtcaaaatac 1500  
 50 ccccaggggg tgctgtagcc aaagtagtgg ttgtcgttgc tggctcctga ttggctggag 1560  
 atttgcttgt agaggtggtt gttgtatgtg ggcaggggccc aggttcgggt gctggtggtg 1620  
 atgactctgt cggccagcca ttgggaatcg caatgccaat ttcttgagga attaccact 1680  
 55 ccatcggcac cctcgttatt gtctgccatt ggtgcgccac tgctgtagc cattgtagta 1740

# EP 1 310 571 A2

	gatcccagac cagagggggc tgctggtggc tgtccgagag gctggggggtc aggtacggag	1800
	tctgcgctctc cagtctgacc aaaattttaat ctttttcttg caggctgctg gcccgttttt	1860
5	ccggttcccg aggaggagtc tggctccaca ggagagtgtc ctaccggcct cttttttccc	1920
	ggagccgtct taacaggctc ctcaaccagg cccagagggt caagaaccct ctttttcgcc	1980
	tggaagactg ctctgccgag gttgccccca aaagacgtat cttctttaag gcgtcctga	2040
10	aactctgctg cggcgtggtt gtacttgagg tacgggttgt ctccgtgtc gagctgccg	2100
	tcgtaggcct tgtcgtgtc gagggccgcg gcgtctgcct cgttgaccgg cttccccttg	2160
	tcgagtcctg tgaagggtcc gaggtacttg taccaggaa gcacaagacc cctgctgtcg	2220
15	tccttatgcc gctctgcggg ctttggtggt ggtgggccag gtttgagctt ccaccactgt	2280
	cttattcctt cagagagagt gtcctcgagc caatctggaa gataaccatc ggcagccata	2340
	cctgatttaa atcattttatt gttcagagat gcagtcaccc aaatccacat tgaccagatc	2400
20	gcaggcagtg caagcgtctg gcacctttcc catgatatga tgaatgtagc acagtttctg	2460
	atacgccttt ttgacgacag aaacgggttg agattctgac acgggaaagc actctaaaca	2520
	gtctttctgt ccgtgagtga agcagatatt tgaattctga ttcattctct cgcattgtct	2580
25	gcagggaaac agcatcagat tcatgccac gtgacgagaa catttgtttt ggtacctgtc	2640
	cgcgtagtgt atcgaagctt ccgcgtctga cgtcgatggc tgcgcaactg actcgcgcgc	2700
	ccgtttgggc tcaattatat ctgcgtcact gggggcgggt cttttcttag ctccaccctt	2760
30	tttgacgtag aattcatgct ccacctcaac cactgatcc tttgccacc ggaaaaagtc	2820
	tttcacttcc tgcttggtga cctttccaaa gtcatgatcc agacggcggg taagttcaaa	2880
	tttgaacatc cggctcttgca acggctgctg gtgctcgaag gtcgttgagt tcccgtcaat	2940
35	cacggcgcac atgttggtgt tggagggtgac gatcacggga gtcgggtcta tctgggccga	3000
	ggacttgcat ttctggtcca cacgcacctt gcttcctcca agaattggctt tggccgactc	3060
	cacgaccttg gcggtcatct tcccctctc ccaccagatc accatcttgt cgacgcaatg	3120
40	gtaaaaggaa agttctcatt gg	3142
	<p>&lt;210&gt; 26            &lt;211&gt; 3075            &lt;212&gt; DNA            &lt;213&gt; new AAV serotype, clone H2</p>	
45		
	<p>&lt;400&gt; 26</p>	
	tgagaacttt cttttcaacg attgcgtcgg acaagatggt gatctggtgg gaggagggga	60
50	agatgaccgc caaggtcgtg gagtcggcca aagccattct tggaggaagc aaggtgcgtg	120
	tggaccagaa atgcaagtcc tcggcccaga tagaccgcac tcccgtgatc gtcacctcca	180
	acaccaacat gtgcgccgtg attgacggga actcaacgac cttcgagcac cagcagccgt	240
55	tgcaagaccg gatgttcaaa tttgaactta cccgccgtct ggatcatgac tttggaaagg	300

EP 1 310 571 A2

	tcaccaagca ggaagtgaag gactttttcc ggtgggcaaa ggatcacgtg gttgaggtgg	360
5	agcatgaatt ctacgtcaaa aagggtggag ctaagaaaag acccgcccc agtgacgcag	420
	atataagtga gcccaaacgg gcgcgcgagt cagttgcgca gccatcaacg tcagacgcgg	480
	aagcttcgat caactacgg gacagggtacc aaaaacaaat gttctcgtca cgtgggcatg	540
10	aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc aaatatctgc	600
	ttcactcacg gacagaaaga ctgttttagag tgctttcccg tgtcagaatc tcaaccctgt	660
	tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat gggaaagggtg	720
15	ccagacgctt gactgcctg cgatctggtc aatgtggatt tggatgactg catctctgaa	780
	caataaatga tttaaatcag gtatggctgc cgatggttat cctccagatt ggctcgagga	840
	cactctctct gaagggataa gacagtgggtg gaagctcaaa cctggcccac caccacaaa	900
20	gcccgcagag cggcataagg acgacagcag gggctctgtg cttcctgggt acaagtacct	960
	cggacccttc aacggactcg acaaggggga gccggtcaac gaggcagacg ccgcgccct	1020
	cgagcacgac aaggcctacg accggcagct cgacagcgga gacaaccgt acctcaagta	1080
25	caaccacgcc gacgcagagt ttcaggagcg ccttaaagaa gatacgtctt ttgggggcaa	1140
	cctcggaaga gcagtcttc aggcgaaaaa gaggggtctt gaacctctgg gcctggttga	1200
	ggaacctgtt aagacggctc cgggaaaaaa gaggcgggtg gagcactctc ctgtggagcc	1260
30	agactcctcc tcgggaaccg gaaaagcggg ccagcggcct gcaagaaaaa gattaaattt	1320
	tggtcagact ggagacgcag actccgtacc tgaccccag cctctcggac agccaccagc	1380
	agccccctct ggtctgggat ctactacaat ggctacaggc agtggcgcac caatggcaga	1440
35	caataacgag ggtgccgatg gagtgggtaa ttcctcagga aattggcatt gcgattccca	1500
	atggctgggc gacagagtca tcaccaccag caccggaacc tgggccctgc ccacatacaa	1560
	caaccacctc tacaagcaaa tctccagcca atcaggagcc agcaacgaca accactactt	1620
40	tggttacagc accccctggg ggtattttga cttcaacaga ttccactgcc acttttcacc	1680
	acgtgactgg caaagactca tcaacaacaa ctggggattc cggcccaaaa gactcaactt	1740
	caagctcttt aatattcaag tcaaagaggt cagcgagaat gacggtacga cgacgattgc	1800
45	caataacctt accagcacgg ttcagggtgt tactgactcg gactaccagc tcccgtacgt	1860
	cctgggctcg gcgcatacag gatgcctccc gccgtttcca gcggacgtct tcatggctcc	1920
	acagtatgga tacctcacc tgaacaacgg gagtcaggcg gtaggacgct cttcctttta	1980
50	ctgcctggag tactttcctt ctcatatgct gcgtactgga aacaactttc agttcagcta	2040
	cacttttgaa gacgtgcctt tccacagcag ctacgctcac agccagagtc tggatcggct	2100
	gatgaatcct ctgatcgacc agtacctgta ttatctgaac aagacacaaa caaatagtgg	2160
55	aactcttcag cagtctcggc tactgttttag ccaagctgga ccaaccaaca tgtctcttca	2220

# EP 1 310 571 A2

	agctaaaaac	tggtgcctg	gaccttgcta	cagacagcag	cgtctgtcaa	aacaggcaaa	2280
5	cgacaacaac	aacagcaact	ttccctggac	tgcagctaca	aagtatcatc	taaatggccg	2340
	ggactcgttg	gttaatccag	gaccagctat	ggccagtcac	aaggatgacg	aagaaaagtt	2400
	tttcccatg	catggaaccc	tgatatttgg	taaaacagga	acaaatgcc	acgacgcgga	2460
10	tttgaaaaat	gtcatgatta	cagatgaaga	agaaatcagg	gccaccaatc	ccgtggctac	2520
	ggagcagtac	gggactgtgt	caaataat	gcaaaactca	aacactggtc	caactactgg	2580
	aactgtcaat	cgccaaggag	cgttacctgg	tatgggtgtg	caggatcgag	acgtgtacct	2640
15	gcagggaccc	atttgggcca	agattcctca	caccgatgga	cactttcatc	cttctccact	2700
	gatgggaggt	tttgactca	aacacccgcc	tcctcagatc	atgatcaaaa	acactcccg	2760
	tccagccaat	cctcccacaa	acttcagttc	tgccaagttt	gcttctttca	tcacacagta	2820
20	ttccacggga	caggtcagcg	tgagatcga	gtgggagctg	cagaaggaga	acagcaaacg	2880
	ctggaatccc	gaaattcagt	acacttccaa	ctacaacaag	tctgttaatg	tggaactttac	2940
	tgtggacact	aatgggtgtg	attcagagcc	tcgccccatt	ggcaccagat	acctgactcg	3000
25	taatctgtaa	ttgcttggtt	atcaataaac	cgtttaattc	gtttcagttg	aactttggtc	3060
	tctgcgaagg	gcgaa					3075
30	<210> 27						
	<211> 3128						
	<212> DNA						
	<213> new AAV serotype, clone 42.8						
	<400> 27						
35	gaattcgccc	tttctacggc	tgcgtaact	ggaccaatga	gaactttccc	ttcaacgatt	60
	gcgtcgacaa	gatgggtgatc	tggtgggagg	agggcaagat	gacggccaag	gtcgtggagt	120
	ccgccaaggc	cattctcggc	ggcagcaagg	tgcgcggtga	ccaaaagtgc	aagtcttccg	180
40	cccagatcga	tcccaccccc	gtgatcgtca	cttccaacac	caacatgtgc	gccgtgattg	240
	acgggaacag	caccaccttc	gagcaccagc	agccgttaca	agaccggatg	ttcaaatttg	300
	aactcaccgc	ccgtctggag	cacgactttg	gcaagggtgac	aaagcaggaa	gtcaaagagt	360
45	tcttccgctg	ggcgcaggat	cacgtgaccg	aggtggcgca	tgagttctac	gtcagaaagg	420
	gtggagccaa	caagagaccc	gcccccgatg	acgcggataa	aagcgagccc	aagcgggcct	480
	gcccctcagt	cgcggatcca	tcgacgtcag	acgcggaagg	agctccggtg	gactttgccg	540
50	acaggtacca	aaacaaatgt	tctcgtcacg	cgggcatgct	tcagatgctg	tttccctgca	600
	agacatgcga	gagaatgaat	cagaatttca	acatttgctt	cacgcacggg	accagagact	660
	gttcagaatg	tttccccggc	gtgtcagaat	ctcaaccgg	cgtcagaaag	aggacgtatc	720
55	ggaaactctg	tgccattcat	catctgctag	ggcgggctcc	cgagattgct	tgctcggcct	780

EP 1 310 571 A2

	gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca	840
5	ggtatggctg ccgatggtta tcttcagat tggctcgagg acaacctctc tgagggcatt	900
	cgcgagtggg gggacttgaa acctggagcc ccgaaacca aagccaacca gcaaaagcag	960
	gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc	1020
10	gacaaggggg agcccgtaa cgcggcgagc gcagcgccc tcgagcacga caaggcctac	1080
	gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag	1140
	tttcaggagc gtctgcaaga agatactgt tttgggggca acctcgggcg agcagtcttc	1200
15	caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct	1260
	cctggaaaga agagaccggt agagccatca cccagcggt ctccagactc ctctacgggc	1320
	atcggcaaga caggccagca gcccgcaaa aagagactca actttgggca gactggcgac	1380
20	tcagagtcag tgcccgacct tcaaccaatc ggagaacccc ccgcaggccc ctctggtctg	1440
	ggatctggtg caatggctgc aggcgggtggc gctccaatgg cagacaataa cgaaggcgcc	1500
	gacggagtgg gtagttcctc aggaaattgg cattgcgatt ccacatggct gggcgacaga	1560
25	gtcatcacca ccagcaccgc aacctgggcc ctccccacct acaacaacca cctctacaag	1620
	caaatctcca acgggacatc gggaggaagc accaacgaca acacctactt cggctacagc	1680
	acccctggg ggtattttga ctttaacaga ttccactgcc acttctcacc acgtgactgg	1740
30	cagcgactca tcaacaacaa ctggggattc cggcccaaga gactcaactt caagctcttc	1800
	aacatccagg tcaaggaggt cagcgagaat gaaggcacca agaccatgc caataacctt	1860
	accagcacga ttcaggtctt tacggactcg gaataccagc tcccgtaagt cctcggctct	1920
35	gcgcaccagg gctgcctgcc tccgttcccg gcggacgtct tcatgattcc tcagtacggg	1980
	tacctgactc tgaacaacgg cagtcaggcc gtgggcccgt cctccttcta ctgcctggag	2040
	tactttcctt ctcaaagtct gagaacgggc aacaactttg agttcagcta ccagtttgag	2100
40	gacgtgcctt ttcacagcag ctacgcgcac agccaaagcc tggaccggct gatgaacccc	2160
	ctcatcgacc agtacctgta ctacctgtct cggactcagt ccacgggagg taccgcagga	2220
	actcagcagt tgctattttc tcaggccggg cctaataaca tgtcggctca ggccaaaaac	2280
45	tggctacccg ggccctgcta ccggcagcaa cgcgtctcca cgacactgtc gcaaaataac	2340
	aacagcaact ttgcttgga cggtgccacc aagtatcatc tgaatggcag agactctctg	2400
	gtaaatcccg gtgtcgctat ggcaacgcac aaggacgacg aagagcgatt ttttccatcc	2460
50	agcggagtct tgatgttttg gaaacaggga gctggaaaag acaacgtgga ctatagcagc	2520
	gttatgctaa ccagtgagga agaaatcaaa accaccaacc cagtggccac agaacagtac	2580
	ggcgtggtgg ccgataacct gcaacagcaa aacgccgctc ctattgtagg ggccgtcaac	2640
55	agtcaaggag ccttacctgg catggtctgg cagaaccggg acgtgtacct gcagggtcct	2700

# EP 1 310 571 A2

	atctggggcca agattcctca cacggacggc aactttcatc cttcgccgct gatgggaggc	2760
	tttggactga aacacccgcc tcctcagatc ctgattaaga atacacctgt tcccgcggat	2820
5	cctccaacta ccttcagtca agccaagctg gcgtcgttca tcacgcagta cagcaccgga	2880
	caggtcagcg tggaaattga atgggagctg cagaaagaga acagcaagcg ctggaaccca	2940
	gagattcagt atacttccaa ctactacaaa tctacaaatg tggactttgc tgtcaatact	3000
10	gaggggtactt attcagagcc tcgccccatt ggcacccggt acctcaccgg taacctgtaa	3060
	ttgcctgtta atcaataaac cggctaattc gtttcagttg aactttggtc tctgcgaagg	3120
	gcgaattc	3128
15		
	<210> 28	
	<211> 3128	
	<212> DNA	
	<213> new AAV serotype, clone 42.15	
20		
	<400> 28	
	gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt	60
	gcgtcgacaa gatggtgatc tgggtggagg agggcaagat gacggccaag gtcgtggagt	120
25	ccgccaaggc cattctcggc ggagcaagg tgcgcgtgga ccaaaagtgc aagtcgtccg	180
	cccagatcga cccaccccc gtgatcgtca cctccaacac caacatgtgc gccgtgattg	240
	acgggaacag caccaccttc gagcaccagc agccgttgca ggaccggatg ttcaaatttg	300
30	aactcaccgg ccgtctggag catgactttg gcaagggtgac aaagcaggaa gtcaaagagt	360
	tcttcgcgtg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg	420
	gtggagccaa caagagacct gccccgatg acgcggataa aagcgagccc aagcgggcct	480
35	gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg	540
	acaggtagca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca	600
	agacatgca gagaatgaat cagaatttca acatttgctt cacgcgcggg accagagact	660
40	gttcagaatg tttccggggc gtgtcagaat ctcaaccggg cgtcagaaag aggacgtatc	720
	ggaaactctg tgccattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct	780
	gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca	840
45	ggataggctg ccgatggtta tcttcagat tggtctgagg acaacctctc tgagggcatt	900
	cgcgagtggt gggacttgaa acctggagcc ccgaaacca aagccaacca gcaaaagcag	960
	gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc	1020
50	gacaaggggg agcccgtaaa cgcggcggac gcagcggccc tcgagcacga caaggcctac	1080
	gaccagcagc tcaaagcggg tgacaatccg tacctgcggg ataaccacgc cgacgccgag	1140
	tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcgggag agcagtcttc	1200
55	caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct	1260

# EP 1 310 571 A2

	cctggaaaga agagaccggt agagccatca cccagcggtt ctccagactc ctctacgggc	1320
5	atcggcaaga caggccagca gcccgcgaaa aagagactca actttgggca gactggcgac	1380
	tcagagtcag tgcccgaccc tcaaccaatc ggagaacccc ccgcaggccc ctctggtctg	1440
	ggatctggta caatggctgc agggcgtggc gctccaatgg cagacaataa cgaaggcgcc	1500
10	gacggagtgg gtagttcctc aggaaattgg cattgcgatt ccacatggct gggcgacaga	1560
	gtcatcacca ccagcacccg aacctgggcc ctccccacct acaacaacca cctctacaag	1620
	caaatctcca acgggacatc gggaggaagc accaacgaca acacctactt cggctacagc	1680
15	acccctggg ggtattttga ctttaacaga ttccactgcc acttctcacc acgtgactgg	1740
	cagcgactca tcaacaacaa ctggggattc cggcccaaga gactcaactt caagctcttc	1800
	aacatccagg tcaaggaggt cacgcagaat gaaggcacca agaccatcgc caataacctt	1860
20	accagcacga ttcagggtctt tacggactcg gaataccagc tcccgtacgt cctcggctct	1920
	gcgccaccag gctgcccgcc tccgttcccg gcggacgtct tcatgattcc tcagtacggg	1980
	tacctgactc tgaacaacgg cagtcaggcc gtgggcggtt cctccttcta ctgcctggag	2040
25	tactttcctt ctcaaattcg gagaacgggc aacaactttg agttcagcta ccagtttgag	2100
	gacgtgcctt ttcacagcag ctacgcgcac agccaaagcc tggaccggct gatgaacccc	2160
	ctcatcgacc agtacctgta ctacctgtct cggactcagt ccacgggagg taccgcagga	2220
30	actcagcagt tgctattttc tcaggccggg cctaataaca tgtcggctca ggccaaaac	2280
	tggctaccgg ggccctgcta ccggcagcaa cgctctcca cgacactgtc gcaaaataac	2340
	aacagcaact ttgcttgga cggtgccacc aagtatcatc tgaatggcag agactctctg	2400
35	gtaaatcccg gtgtcgctat ggcaacgcac aaggacgacg aagagcgatt tttccatcc	2460
	agcggagtct tgatgtttgg gaaacaggga gctggaaaag acaacgtgga ctatagcagc	2520
	gttatgctaa ccagtgagga agaaatcaaa accaccaacc cagtggccac agaacagtac	2580
40	ggcgtggtgg ccgataacct gcaacagcaa aacgcgcgtc ctattgtagg ggccgtcaac	2640
	agtcaaggag ccttacctgg catggtctgg cagaaccggg acgtgtacct gcagggtcct	2700
	atctgggcca agattcctca cacggacggc aactttcatc cttcgccgct gatgggaggc	2760
45	tttggactga aacaccgcc tccctcagatc ctgattaaga atacacctgt tcccgcggat	2820
	cctccaacta ccttcagtca agccaagctg gcgtcgttca tcacgcagta cagcaccgga	2880
	caggtcagcg tggaaattga atgggagctg cagaaagaga acagcaagcg ctggaacca	2940
50	gagattcagt atacttcaa ctactacaaa tctacaaatg tggactttgc tgtcaatact	3000
	gagggtactt attcagagcc tcgccccatt ggcaccggtt acctcaccg taacctgtaa	3060
	ttgcctgtta atcaataaac cggttaattc gtttcagttg aactttggtc tctgcgaagg	3120
55	gcgaattc	3128

# EP 1 310 571 A2

<210> 29  
 <211> 3197  
 <212> DNA  
 5 <213> new AAV serotype. clone 42.5b

<400> 29  
 gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60  
 gcgctcgacaa gatggtgatc tgggtggagg agggcaagat gacggccaag gtcgtggagt 120  
 10 ccgccaaggc cattctcggc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcgtccg 180  
 ccagatcga cccaccccc gtgatcgtca cctccaacac caacatgtgc gccgtgattg 240  
 acgggaacag caccaccttc gagcaccagc agccgttaca agaccggatg ttcaaatttg 300  
 15 aactcaccgc ccgtctggag cagcactttg gcaagggtgac aaagcaggaa gtcaaagagt 360  
 tcttcgcgtg ggcgaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg 420  
 gtggagccaa caagagaccc gccccgatg acgcggataa aagcgagccc aagcgggcct 480  
 20 gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540  
 acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca 600  
 25 agacatgcga gagaatgaat cagaatttca acatttgctt cagcacggg accagagact 660  
 gttcagaatg tttccccggc gtgtcagaat ctcaaccggg cgtcagaaaag aggacgtatc 720  
 ggaaactctg tgccattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct 780  
 30 gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca 840  
 ggtatggctg ccgatgggta tcttcagat tggctcgagg acaacctctc tgagggcatt 900  
 cgcgagtggg gggacttgaa acctggagcc ccgaaacca aagccaacca gcaaaagcag 960  
 35 gacgacggcc ggggtctggt gcttcttggc tacaagtacc tcggaccctt caacggactc 1020  
 gacaagggag agccgggtcaa cgaggcagac gccgcggccc tcgagcacga caaggcctac 1080  
 gacaagcagc tcgagcaggg ggacaacccg tacctcaagt acaaccacgc cgacgccgag 1140  
 40 tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggag agcagtcttc 1200  
 caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260  
 cctggaaaga agagaccggg agagccatca cccagcgtt ctccagactc ctctacgggc 1320  
 45 atcggaaga caggccagca gccgcgaaa aagagactca actttgggca gactggcgac 1380  
 tcagagtcag tgcccgacct tcaaccaatc ggagaacccc ccgcaggccc ctctggtctg 1440  
 ggatctggtg caatggctgc agcggtggc gctccaatg cagacaataa cgaaggcgcc 1500  
 50 gacggagtgg gtagttcttc aggaaatttg cattgcgatt ccacatggct gggcgacaga 1560  
 gtcatcacca ccagcaccgc aacctgggac ctccccacct acaacaacca cctctacaag 1620  
 caaatctcca acgggacatc gggaggaagc accaaccgac acacctactt cggctacagc 1680  
 55 accccctggg ggtattttga cttaacaga ttccactgcc acttctcacc acgtgactgg 1740

# EP 1 310 571 A2

	cagcgactca tcaacaacaa ctggggattc cggcccaaga gactcaactt caagctcttc	1800
	aacatccagg tcaaggaggt cacgcagaat gaaggcacca agaccatcgc caataacctt	1860
5	accagcacga ttcagggtctt tacggactcg gaataccagc tcccgtacgt cctcggctct	1920
	gcgcaccagg gctgcctgcc tccgttcccg gcggacgtct tcatgattcc tcagtacggg	1980
10	tacctgactc tgaacaacgg cagtcaggcc gtgggccgtt cctccttcta ctgcctggag	2040
	tactttcctt ctcaaatgct gagaacgggc aacaactttg agttcagcta ccagtttgag	2100
	gacgtgcctt ttcacagcag ctacgcgcac agccaaagcc tggaccggct gatgaacccc	2160
15	ctcatcgacc agtacctgta ctacctgtct cggactcagt ccacgggagg taccgcagga	2220
	actcagcagt tgctattttc tcaggccggg cctaataaca tgtcggctca ggccaaaaac	2280
	tggctacccg ggccctgcta ccggcagcaa cgcgtctcca cgacactgtc gcaaaataac	2340
20	aacagcaact ttgcttggac cggtgccacc aagtatcatc tgaatggcag agactctctg	2400
	gtaaatcccg gtgtcgctat ggcaacgcac aaggacgacg aagagcgatt ttttccatcc	2460
	agcggagtct tgatgtttgg gaaacaggga gctggaaaag acaacgtgga ctatagcagc	2520
25	gttatgctaa ccagtgagga agaaatcaaa accaccaacc cagtggccac agaacagtac	2580
	ggcgtggtgg ccgataacct gcaacagcaa aacgccgctc ctattgtagg ggccgtcaac	2640
	agtcaaggag ccttacctgg catgggtctgg cagaaccggg acgtgtacct gcagggctct	2700
30	atctgggcca agattcctca cacggacggc aactttcatc cttcgcgct gatgggaggc	2760
	tttggactga aacaccgcc tcctcagatc ctgattaaga atacacctgt tccgcggat	2820
	cctccaacta ccttcagtca agccaagctg gcgtcggtca tcacgcagta cagcaccgga	2880
35	caggtcagcg tggaaattga atgggagctg cagaaagaga acagcaagcg ctggaaccca	2940
	gagattcagt atacttccaa ctactacaaa tctacaaatg tggactttgc tgtcaatact	3000
	gaggggtactt attcagagcc tcgccccatt ggcaccggtt acctcaccg taacctgtaa	3060
40	ttgcctgtta atcaataaac cggttaattc gtttcagttg aactttggtc tctgcgaagg	3120
	gcgaattcgt ttaaacctgc aggactagtc cttttagtga gggttaattc tgagcttggc	3180
	gtaatcatgg gtcatag	3197
45	<p>&lt;210&gt; 30            &lt;211&gt; 2501            &lt;212&gt; DNA            &lt;213&gt; new AAV serotype, clone 42.1b</p>	
50	<400> 30 gaattcgccc ttggctgcgt caactggacc aatgagaact ttcccttcaa cgattgcgtc	60
	gacaagatgg tgatctggtg ggaggagggc aagatgacgg ccaaggctcg ggagtccgcc	120
55	aaggccattc atcatctgct ggggcgggct cccgagattg cttgctcggc ctgcgatctg	180
	gtcaacgtgg acctggatga ctgtgtttct gagcaataaa tgacttaaac caggtatggc	240

EP 1 310 571 A2

	tgccgatggt tatcttccag attggctcga ggacaacctc tctgagggca ttcgcgagtg	300
	gtgggacttg agacctggag ccccgaaacc caaagccaac cagcaaaagc aggacgacgg	360
5	ccggggctctg gtgcttcttg gctacaagta cctcggacct ttcaacggac tcgacaaggg	420
	agagccggtc aacgaggcag acgccgcggc cctcgagcac gacaaggcct acgacaagca	480
	gctcgagcag ggggacaacc cgtacctcaa gtacaaccac gccgacgccg agtttcagga	540
10	gcgtcttcaa gaagatacgt cttttggggg caacctcggg cgagcagtct tccaggccaa	600
	gaagcggggt ctcgaacctc tcggctctgt tgaggaaggc gctaagacgg ctccctggaaa	660
	gaagagacct atagaatccc ccgactcctc cacgggcacg ggcaagaaaag gccagcagcc	720
15	cgctaaaaag agactcaact ttgggcagac tggcgactca gagtcaagtgc ccgaccctca	780
	accaatcgga gaaccccccg caggccccctc tggctctggga tctggcaciaa tggctgcagg	840
	cggtggcgct ccaatggcag acaataacga aggcgcgcgac ggagtgggta gttcctcagg	900
20	aaattggcat tgcgattcca catggctggg cgacagagtc atcaccacca gcacccgaac	960
	ctggggccctc cccacctaca acaaccacct ctacaagcaa atctccaacg ggacatcggg	1020
	aggaagcacc aacgacaaca cctacttcgg ctacagcacc ccctgggggt attttgactt	1080
25	taacagattc cactgccact tctcaccacg tgactggcag cgactcatca acaacaactg	1140
	gggattccgg cccaagagac tcaacttcaa gctcttcaac atccaggta aggaggtcac	1200
	gcagaatgaa ggcaccaaga ccatcgccaa taaccttacc agcacgattc aggtctttac	1260
30	ggactcggaa taccagctcc cgtacgtcct cggctctgcg caccagggtc gcctgcctcc	1320
	gttcccggcg gacgtcttca tgattcctca gtacgggtac ctgactctga acaacggcag	1380
	tcaggccgtg ggccgttctt ccttctactg cctggagtac tttccttctc aaatgctgag	1440
35	aacgggcaac aactttgagt tcagctacca gtttgaggac gtgccttttc acagcagcta	1500
	tgcgcacagc caaagcctgg accggctgat gaacccccctc atcgaccagt acctgtacta	1560
	cctgtctcgg actcagtcca cgggagggtac cgcaggaact cagcagttgc tattttctca	1620
40	ggccgggcct aataacatgt cggctcaggc caaaaactgg ctaccgggc cctgctaccg	1680
	gcagcaacgc gtctccacga cagtgtcgca aaataacaac agcaactttg cttggaccgg	1740
	tgccaccaag tatcatctga atggcagaga ctctctggta aatcccgggt tcgctatggc	1800
45	aacgcacaag ggcgacgaag agcgattttt tccatccagc ggagtcttga tgtttgggaa	1860
	acagggagct gaaaaagaca acgtagacta tagcagcggt atgctaacca gtgaggaaga	1920
	aatcaaaacc accaaccacg tggccacaga acagtacggc gtggtggccg ataacctgca	1980
50	acagcaaaac gccgctccta ttgtaggggc cgtcaacagt caaggagcct tacctggcat	2040
	ggtctggcag aaccgggacg tgtacctgca gggctctatc tgggccaaga ttcttcacac	2100
55	ggacggcaac tttcatcctt cgccgctgat gggaggcttt ggactgaaac acccgctctc	2160

# EP 1 310 571 A2

	tcagatcctg attaagaata cacctgttcc cgcggatcct ccaactacct tcagtcaagc	2220
	caagctggcg tcgttcatca cgcagtagag caccggacag gtcagcgtgg aaattgaatg	2280
5	ggagctgcag aaagagaaca gcaagcgtg gaaccagag attcagtata cttccaacta	2340
	ctacaaatct acaaagtgg actttgctgt caatactgag ggtacttatt cagagcctcg	2400
	ccccattggc acccgttacc tcaccgtaa cctgtaattg cctgttaatc aataaaccgg	2460
10	ttgattcgtt tcagttgaac tttgggtctca agggcgaatt c	2501
	<210> 31	
	<211> 3113	
15	<212> DNA	
	<213> new AAV serotype, clone 42.13	
	<400> 31	
	gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt	60
20	gcgctcgacaa gatgggtgatc tgggtgggagg agggcaagat gacggccaag gtcgtggagt	120
	ccgccaaggc cattctcggc ggcagcaagg tgcgctgga ccaaaagtgc aagtcgtccg	180
	cccagatcga tcccaccccc gtgatcgtca cttccaacac caacatgtgc gccgtgattg	240
25	acgggaacag caccaccttc gagcaccagc agccgttaca agaccggatg ttcaaatttg	300
	aactcaccgc ccgtctggag catgactttg gcaagggtgac aaagcaggaa gtcaaagagt	360
	tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg	420
30	gtggagccaa caagagaccg gccccgatg acgcgataa aagcgagccc aagcgggcct	480
	gccccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg	540
	acaggtagca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca	600
35	agacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg accagagact	660
	gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc	720
	ggaaactctg tgccattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct	780
40	gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca	840
	ggtatggctg ccgatggtta tcttccagat tggctcgagg acaacctctc tgagggcatt	900
	cgcgagtggg gggacttgaa acctggagcc ccgaaacca aagccaacca gcaaaagcag	960
45	gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc	1020
	gacaaggggg agcccgtaa cgcggcggac gcagcggccc tcgagcacga caaggcctac	1080
	gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag	1140
50	tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc	1200
	caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct	1260
	cctggaaaga agagacccat agaatcccc gactcctcca cgggcacgag caagaaaggc	1320
55	cagcagcccc ctaaaaagaa gctcaacttt gggcagactg gcgactcaga gtcagtgcc	1380

# EP 1 310 571 A2

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50

gaccctcaac caatcggaga acccccgcga ggccctctctg gtctgggac tggtaacaatg 1440  
 gctgcaggcg gtggcgctcc aatggcagac aataacgaag gcgccgacgg agtgggtagt 1500  
 tcctcaggaa attggcattg cgattccaca tggctgggcg acagagtcac caccaccagc 1560  
 acccgaacct gggccctccc cacctacaac aaccacctct acaagcaaatt ctccaacggg 1620  
 acatcgggag gaagcaccaa cgacaacacc tacttcgggt acagcacccc ctgggggtat 1680  
 tttgacttta acagattcca ctgccacttc tcaccacgtg actggcagcg actcatcaac 1740  
 aacaactggg gattccggcc caagagactc aacttcaagc tcttcaacat ccagggtcaag 1800  
 gaggtcacgc agaataagg caccaagacc atcgccaata accttaccag cagcattcag 1860  
 gtctttacgg actcggata ccagctcccg tacgtcctcg gctctgcgca ccagggtgc 1920  
 ctgcctccgt tcccggcgga cgtcttcattg attcctcagt acgggtacct gactctgaac 1980  
 aacggcagtc aggcgtggg ccgttctctc ttctactgcc tggagtactt tcctttctca 2040  
 atgctgagaa cgggcaacaa ctttgagttc agctaccagt ttgaggacgt gccttttctac 2100  
 agcagctatg cgcacagcca aagcctggac cggctgatga acccctcat cgaccagtac 2160  
 ctgtactacc tgtctcggac tcagtcacag ggaggtaccg caggaaactca gcagttgcta 2220  
 ttttctcagg ccgggcctaa taacatgtcg gctcaggcca aaaactggct acccgggccc 2280  
 tgctaccggc agcaacgcgt ctccacgaca gtgtcgcaaa ataacaacag caactttgct 2340  
 tggaccggtg ccaccaagta tcatctgaat ggcagagact ctctggtaaa tcccgggtgc 2400  
 gctatggcaa cgcacaaggg cgacgaagag cgattttttc catccagcgg agtcttgatg 2460  
 tttgggaaac agggagctgg aaaagacaac gtggactata gcagcgttat gctaaccagt 2520  
 gaggaagaaa tcaaaaccac caaccagtg gccacagaac agtacggcgt ggtggccgat 2580  
 aacctgcaac agcaaaacgc cgctcctatt gtaggggccc tcaacagtca aggagcctta 2640  
 cctggcatgg tctggcagaa ccgggacgtg tacctgcagg gtcctatctg ggccaagatt 2700  
 cctcacacgg acggcaactt tcctccttcg ccgtgatgg gaggccttgg actgaaacac 2760  
 ccgcctctc agatcctgat taagaataca cctgttcccg cggatcctcc aactaccttc 2820  
 agtcaagcca agctggcgtc gttcatcacg cagtacagca ccggacaggt cagcgtggaa 2880  
 attgaatggg agctgcagaa agagaacagc aagcgtgga acccagagat tcagtatact 2940  
 tccaactact acaaactctac aaatgtggac tttgctgtca atactgaggg tacttattca 3000  
 gagcctcgcc ccattggcac ccgttacctc acccgtagcc tgtaattgcc tgtaaatcaa 3060  
 taaaccgggt gattcgtttc agttgaactt tggctctctgc gaaggcgaa ttc 3113

<210> 32  
 <211> 3113  
 <212> DNA  
 <213> new AAV serotype, clone 42.3a

# EP 1 310 571 A2

<400> 32  
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60  
5 gcgtcgacaa gatggtgac tggtgggagg agggcaagat gacggccaag gtcgtggagt 120  
ccgccaaggc cattctcggc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcgtccg 180  
cccagatcga tcccaccccc gtgatcgtca cttccaacac caacatgtgc gccgtgattg 240  
10 acgggaacag caccacctc gagcaccagc agccgttaca agaccggatg ttcaaatttg 300  
aactcaccgg ccgtctggag catgactttg gcaaggtgac aaagcaggaa gtcaaagagt 360  
tcttccgctg ggcgaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg 420  
15 gtggagccaa caagagaccc gccccgatg acgcgataa aagcgagccc aagcgggcct 480  
gcccctcagt cgcgatcca tcgacgtcag acgcggaag agctccggtg gactttgccg 540  
acaggtagca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg cttccctgca 600  
20 agacatgcca gagaatgaat cagaatttca gcatttgctt cagcacggg accagagact 660  
gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc 720  
ggaaactctg tgccattcat catctgctgg ggcggtctcc cgagattgct tgctcggcct 780  
25 gcgatctggt caacgtggac ctggaagact gtgtttctga gcaataaatg acttaaacca 840  
ggtatggctg ccgatggta tcttcagat tggctcgagg acaacctctc tgagggcatt 900  
cgcgagtggg gggacttgaa acctggagcc ccgaaacca aagccaacca gcaaaagcag 960  
30 gacgacggcc ggggtctggt gcttctggc tacaagtacc tcggaccctt caacggactc 1020  
gacaaggggg agcccgtaa cgcgcgagc gcagcgccc tcgagcacga caaggcctac 1080  
gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag 1140  
35 tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggag agcagtcttc 1200  
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260  
cctggaaaga agagacccat agaatcccc gactcctcca cgggcatcgg caagaaaggc 1320  
40 cagcagcccg ctaaaaagaa gctcaacttt ggcagactg gcgactcaga gtcagtgcc 1380  
gacctcaac caatcggaga accccccgca ggccctctg gtctgggatc tggtaaatg 1440  
gctgcaggcg gtggcgctcc aatggcagac aataacgaag gcgccgacgg agtgggtagt 1500  
45 tcctcaggaa attggcattg cgattccaca tagctgggag acagagtcac caccaccagc 1560  
acctgaacct ggccctccc cacctacaac aaccacctct acaagcaaact ctccaacggg 1620  
acatcgggag gaagcaccaa cgacaacacc tacttcggct acagcaccac ctgggggtat 1680  
50 tttgacttta acagattcca ctgccacttc tcaccacgtg actggcagcg actcatcaac 1740  
aacagctggg gattccggcc caagagactc aacttcaagc tcttcaacat ccaggtcaag 1800  
gaggtcacgc agaataagc caccaagacc atcgccaata accttaccag cagattcag 1860  
55 gtctttacgg actcgaata ccagctcccg tacgtcctcg gctctgcgca ccagggtgct 1920

# EP 1 310 571 A2

	ctgcctccgt tcccggcgga cgtcttcacg attcctcagt acgggtacct gactctgaac	1980
	aacggcagtc aggcgcgtggg ccgttcctcc ttctactgcc tggagtactt tcctttctcaa	2040
5	atgctgagaa cgggcaacaa ctttgagttc agctaccagt ttgaggacgt gccttttcac	2100
	agcagctacg cgcacagcca aagcctggac cggctgatga accccctcat cgaccagtac	2160
	ctgtactacc tgtctcggac tcagtccacg ggaggtaccg caggaactca gcagttgcta	2220
10	ttttctcagg ccgggcctaa taacatgtcg gctcaggcca aaaactggct acccgggccc	2280
	tgctaccggc agcaacgcgt ctccacgaca ctgtcgcaaa ataacaacag caactttgct	2340
	tggaccgggtg ccaccaagta tcatctgaat ggagagact ctctggtaaa tcccgggtg	2400
15	gctatggcaa cgcacaagga cgacgaagag cgattttttc catccagcgg agtcttgatg	2460
	tttgggaaac agggagctgg aaaagacaac gtggactata gcagcgttat gctaaccagt	2520
	gaggaagaaa tcaaaaccac caaccacgtg gccacagaac agtacggcgt ggtggccgat	2580
20	aacctgcaac agcaaaacgc cgctcctatt gtaggggccc tcaacagtca aggagcctta	2640
	cctggcatgg tctggcagaa ccgggacgtg tacctgcagg gtcctatctg ggccaagatt	2700
	cctcacacgg acggcaactt tcatccttcg ccgctgatgg gaggctttgg actgaaacac	2760
25	ccgcctcttc agatcctgat taagaataca cctgttcccg cggatcctcc aactaccttc	2820
	agtcaagcca agctggcgtc gttcatcacg cagtacagca ccggacaggt cagcgtggaa	2880
	attgaatggg agctgcagaa agagaacagc aagcgtgga acccagagat tcagtatact	2940
30	tccaactact acaaactctac aaatgtggac tttgctgtca atactgaggg tactttatca	3000
	gagcctcgcc ccattggcac ccgttacctc acccgtaacc tgtaattgcc tgttaatcaa	3060
35	taaaccgggtt aattcgtttc agttgaactt tgggtctctgc gaagggcgaa ttc	3113
	<210> 33	
	<211> 2504	
	<212> DNA	
40	<213> new AAV serotype, clone 42.4	
	<400> 33	
	gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt	60
	gcgtcgacaa gatgggtgatc tgggtgggagg agggcaagat gacggccaag gtcgtggagt	120
45	ccgccaaggc cattcatcat ctgctggggc gggtcccga gattgcttgc tcggcctgcg	180
	atctggtcaa cgtggacctg gatgactgtg tttctgagca ataaatgact taaaccaggt	240
	atggctgccg atggttatct tccagattgg ctcgaggaca acctctctga gggcattcgc	300
50	gagtgggtggg acttgaaacc tggagccccg aaaccctaaag ccaaccagca aaagcaggac	360
	gacggccggg gtctggtgct tcctggctac aagtacctcg gacccttcaa cggactcgac	420
	aagggagagc cgggtcaacga ggcagacgcc gcggccctcg agcacgacaa ggcctacgac	480
55	aagcagctcg agcaggggga caacccttac ctcaagtaca accacgccga cgccgagttt	540

EP 1 310 571 A2

	caggagcgtc ttcaagaaga tacgtctttt gggggcaacc tcgggcgagc agtctttccag	600
	gccaagaagc gggttctcga acctctcggt ctggttgagg aaggcgctaa gacggctcct	660
5	ggaaagaaga gacccataga atccccgac tcctccacgg gcacgcgcaa gaaaggccag	720
	cagcccgtc aaaagaagct caactttggg cagactggcg actcagagtc agtgcccgc	780
	cctcaaccaa tcggagaacc cccgcaggc ccctctggtc tgggatctgg tacaatggct	840
10	gcaggcgggtg gcgctccaat ggcagacaat aacgaaggcg ccgacggagt gggtaatgcc	900
	tccggaaatt ggcatcgca ttccacatgg ctgggcgaca gagtcacac caccagcacc	960
	cgcacctggg ccctgcccac ctacaacaac cacctctaca agcagatata aagtcagagc	1020
15	ggggctacca acgacaacca cttcttcggc tacagcacc cctggggcta ttttgacttc	1080
	aacagattcc actgccactt ctcatcacgt gactggcagc gactcatcaa caacaactgg	1140
	ggattccggc ccaagagact caacttcaag ctcttcaaca tccaggtaaa ggaggtcacg	1200
20	cagaatgaag gcaccaagac catcgccaat aaccttacca gcacgattca ggtctttacg	1260
	gactcggaat accggctccc gtacgtctc ggctctggc accagggtcg cctgcctccg	1320
	ttcccgcgcg acgtcttcat gattcctcag tacgggtacc tgactctgaa caacggcagt	1380
25	caggccgtgg gccgttcctc cttctactgc ctggagtact ttccttctca aatgctgaga	1440
	acgggcaaca actttgagtt cagctaccag tttagggacg tgccttttca cagcagctac	1500
	gcgcacagcc aaagcctgga ccggctgatg aacccccca tcgaccagta cctgtactac	1560
30	ctgtctcgga ctcatgccac gggaggtaac gcaggaactc agcagttgct attttctcag	1620
	gccgggccta ataacatgtc ggctcaggcc aaaaactggc taccggggcc ctgctaccgg	1680
	cagcaacgcg tctccacgac actgtcgcaa aataacaaca gcaactttgc ttggaccggt	1740
35	gccaccaagt atcatctgaa tggcagagac tctctggtaa atcccggtgt cgctatggca	1800
	acgcacaagg acgacgaaga gcgattttt ccatccagcg gactcttgat gtttgggaaa	1860
	caggagctg gaaaagacaa cgtggactat agcagcgcta tgctaaccag tgaggaagaa	1920
40	atcaaaacca ccaaccaggt ggccacagaa cagtacggcg tggtagccga taacctgcaa	1980
	cagcaaaacg ccgctcctat tgtaggggcc gtcaacagtc aaggagcctt acctggcatg	2040
	gtctggcaga accgggacgt gtacctgcag ggtcctatct gggccaagat tcctcacacg	2100
45	gacggcaact ttcaccttc gccgctgatg ggaggtttg gactgaaaca cccgcctcct	2160
	cagatcctga ttaagaatac acctgttccc gcggtacctc caactacctt cagtcaagcc	2220
	aagccggcgt cgttcatcac gcagtacagc accggacagg tcagcgtgga aattgaatgg	2280
50	gagctgcaga aagagaacag caagcgctgg aaccagaga ttcagtatac ttccaactac	2340
	tacaaatcta caaatgtgga ctttgctgtc aatactgagg gtacttattc agagcctcgc	2400
55	cccattggca ccggttacct caccgtaac ctgtaattgc ctgttaatca ataaaccggt	2460

# EP 1 310 571 A2

	taattcgttt cagttgaact ttggtctctg cgaagggcga attc	2504
5	<210> 34 <211> 3106 <212> DNA <213> new AAV serotype, clone 42.5a	
10	<400> 34 gaattcggccc ttctacggct gcgtcaactg gaccaatgag aactttccct tcaacgattg	60
	cgctgcacaag atggtgatct ggtgggagga gggcaagatg acggccaagg tcgtggagtc	120
	cgccaaggcc attctcggcg gcagcaagggt gcgcgtggac caaaagtgc agtcgtccgc	180
15	ccagatcgac cccacccccg tgatcgtcac ctccaacacc aacatgtgcg ccgtgattga	240
	cgggaaacagc accaccttcg agcaccagca gccgttgag gaccggatgt tcaaatttga	300
	actcaccgcg cgtctggagc atgacttttg caaggcgaca aagcaggaag tcaaagagtt	360
20	cttcgctggt gcgcaggatc acgtgaccga ggtggcgcat gagttctacg tcagaaaggg	420
	tggagccaac aagagacccg ccccgatga cgcgataaa agcgagccca agcgggccc	480
	cccctcagtc gcggatccat cgacgtcaga cgcggaagga gctccggttg actttgccga	540
25	caggtaccaa aacaaatgtt ctcgtcacgc gggcatgctt cagatgctgt ttccctgcaa	600
	aacatgagag agaataatc agaatttcaa catttgcttc acgcacggga ccagagactg	660
	ttcagaatgt ttccccggcg tgtcagaatc tcaaccggtc gtcagaaaga ggacgtatcg	720
30	gaaactctgt gccattcatc atctgctggg gcgggctccc gagattgctt gctcggcctg	780
	cgatctggtc aacgtggacc tggatgactg tgtttctgag caataaatga cttaaaccag	840
	gtatggctgc cgatgggttat cttccagatt ggctcgagga caacctctct gagggcattc	900
35	gcgagtgggt ggacttgaaa cctggagccc cgaaacccaa agccaaccag caaaagcagg	960
	acgacggccg ggggtctggt cttcctggct acaagtacct cggacccttc aacggactcg	1020
	acaagggaga gccggtcaac gaggcagacg ccgcggccct cgagcacgac aaggcctacg	1080
40	acaagcagct cgagcagggg gacaaccgct acctcaagta caaccacgcc gacgccgagt	1140
	ttcaggagcg tcttcaagaa gatacgtctt ttgggggcaa cctcgggcca gcagtcttcc	1200
	gggccaagaa gcgggttctc gaacctctcg gtctggttga ggaaggcgct aagacggctc	1260
45	ctggaaagaa gagaccata gaatcccccg actcctccac gggcatcggc aagaaaggcc	1320
	agcagcccgc taaaaagaag ctcaactttg ggcagactgg cgactcagag tcagtgcccg	1380
	accccccaacc tctcggagaa cctcccgcg cgccctcagg tctgggatct ggtacaatgg	1440
50	ctgcaggcgg tggcgaccca atggcagaca ataacgaagg cgccgacgga gtgggtaatg	1500
	cctccgaaa ttggcattgc gattccacat ggctgggcca cagagtcac accaccagca	1560
	cccgcacctg ggccctgccc acctacaaca accacctcta caagcagata tcaagtcaga	1620
55	gcgggggctac caacgacaac cacttcttcg gctacagcac cccctggggc tattttgact	1680

# EP 1 310 571 A2

	tcaacagatt ccaactgccac ttctcaccac gtgactggca gcgactcatc aacaacaacc	1740
	ggggattccg gccagaaaag ctgcgggttca agttgttcaa catccaggtc aaggagggtca	1800
5	cgacgaacga cggcgttacg accatcgcta ataaccttac cagcacgatt caggtcttct	1860
	cggactcggga gtaccaactg ccgtacgtcc tcggctctgc gcaccagggc tgcctccctc	1920
	cgttccctgc ggacgtgttc atgattcctc agtacggata tctgactcta aacaacggca	1980
10	gtcagtctgt gggacgttcc tccttctact gcctggagta ctttcccttct cagatgctga	2040
	gaacgggcaa taactttgaa ttcagctacc agtttgagga cgtgcccttt cacagcagct	2100
	acgcgcacag ccaaagcctg gaccggctga tgaacccct catcgaccag tacctgtact	2160
15	acctgtctcg gactcagtc acgggaggt cgcaggaac tcagcagttg ctattttctc	2220
	aggccggggc taataacatg tcggctcagg ccaaaaactg gctaccggg ccctgctacc	2280
	ggcagcaacg cgtctccacg acactgtcgc aaaataacaa cagcaacttt gcttggaccg	2340
20	gtgccaccaa gtatcatctg aatggcagag actctctggt aaatcccgtg gtcgctatgg	2400
	caacgcacaa ggacgacgaa gagcgatttt ttccatccag cggagtcttg atgtttggga	2460
	aacaggggagc tggaaaagac aacgtggact atagcagcgt tatgctaacc agtgaggaag	2520
25	aaatcaaaac caccaaccca gtggccacag aacagtacgg cgtggtggcc gataacctgc	2580
	aacagcaaaa cgccgctcct attgtagggg ccgtcaacag tcaaggagcc ttacctggca	2640
	tggcctggca gaaccgggac gtgtacctgc agggctctat ctggggccaag attcctcaca	2700
30	cggacggcaa ctttcatcct tcgccgtga tgggaggctt tggactgaaa caccgcctc	2760
	ctcagatcct gattaagaat acacctgttc ccgcggatcc tccaactacc ttcagtcaag	2820
	ccaagctggc gtcgttcac acgcagtaca gcaccggaca ggtcagcgtg gaaattgaat	2880
35	gggagctgca gaaagagaac agcaagcgt ggaaccaga gattcagtat acttccaact	2940
	actacaaatc taaaaatgtg gactttgctg tcaatactga gggacttat tcagagcctc	3000
	gccccattgg caccggttac ctcaccgta acctgtaatt gcctgttaat caataaaccg	3060
40	gttaattcgt ttcagttgaa ctttggctc tcggaagggc gaattc	3106
	<210> 35	
45	<211> 2489	
	<212> DNA	
	<213> new AAV serotype, clone 42.10	
	<400> 35	
	gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt	60
50	gcgtcgacaa gatgggtgatc tgggtgggagg agggcaagat gacggccaag gtcgtgaagt	120
	ccgccaaggc cattcatcat ctgctggggc gggctcccga gattgcttgc tcggcctgcg	180
	atctggtcaa cgtggacctg gatgactgtg tttctgagca ataatgact taaaccaggt	240
55	atggctgccg atggttatct tccagattgg ctcgaggaca acctctctga gggcattcgc	300

EP 1 310 571 A2

	gagtgggtggg acttgaaacc tggagccccg aaaccctaaag ccaaccagca aaagcaggac	360
	gacggccggg gtctgggtgct tcctggctac aagtacctcg gacccttcaa cggactcgac	420
5	aagggagagc cggtaacga ggcagacgcc gcggccctcg agcacgacaa ggcctacgac	480
	aagcagctcg agcaggggga caaccctgac ctcaagtaca accacgccga cggcagagttt	540
10	caggagcgctc ttcaagaaga tacgtctttt gggggcaacc tcgggagagc agtcttccag	600
	gccagaagc gggttctcga acctctcggt ctgggtgagg aaggcgctaa gacggctcct	660
	ggaaagaaga gacccataga atccccgcac tcctccacgg gcacggcag gaaaggccag	720
15	cagcccgcta aaaagaagct caactttggg cagactggcg actcagagtc agtgcccgac	780
	cctcaaccaa tcggagaacc ccccgaggc cctctgggtc tgggatctgg tacaatggct	840
	gcaggcggtg gcgctccaat ggcagacaat aacgaaggcg ccgacggagt gggtaatgcc	900
20	tccggaaatt ggcattgcga ttccacatgg ctgggagaca gagtcacac caccagcacc	960
	cgcacctggg ccctgccac ctacaacaac cacctctaca agcagatata aagtcagagc	1020
	ggggctacca acgacaacca cttcttcggc tacagcacc cctggggcta ttttgacttc	1080
25	aacagattcc actgccactt ctcaccacgt gactggcagc gactcatcaa caacaactgg	1140
	ggattccggc ccagaaagct gcggttcaag ttgttcaaca tccagggtcaa ggaggtcacg	1200
	acgaacgacg gcgttacgac catcgccaat aaccttacca gcacgattca ggtcttctcg	1260
30	gactcggagt accaactgcc gtacgtcctc ggctctgcgc accagggtcg cctccctccg	1320
	ttccctgcgg acgtgttcat gattcctcag tacggatata tgactctaaa caacggcagt	1380
	cagtctgtgg gacgttcctc cttctactgc ctggagtact ttcttctca gatgctgaga	1440
35	acgggcaata actttgaatt cagctacacc tttgaggaag tgcctttcca cagcagctat	1500
	gcgcacagcc agagcctgga ccggcctgatg aatccctca tcgaccagta cctgtactac	1560
	ctggcccgga cccagagcac tacggggtcc acaaggagc tgcagttcca tcaggctggg	1620
40	cccaacacca tggccgagca atcaaagaac tggctgccc gaccctgtta tcggcagcag	1680
	agactgtcaa aaaacataga cagcaacaac aacagtaact ttgcctggac cggggccact	1740
	aaataccatc tgaatggtag aaattcatta accaaccggg gcgtagccat ggccaccaac	1800
45	aaggacgacg aggaccagtt ctttcccatc aacggagtgc tggtttttgg caaaacgggg	1860
	gctgccaaaca agacaacgct ggaaaacgtg ctaatgacca gcgaggagga gatcaaaacc	1920
	accaatcccg tggctacaga agaatacggg gtggtctcca gcaacctgca atcgtctacg	1980
50	gccggacccc agacacagac tgtcaacagc cagggggctc tgcccgcat ggtctggcag	2040
	aaccgggacg tgtacctgca ggggtccatc tggggcaaaa ttctcacac ggacggcaac	2100
	tttaccctgt cttccctgat gggcggtttt ggactcaaac acccgctcc tcaaattctc	2160
55	atcaaaaaca ccccggtacc tgctaatact ccagaggtgt ttactcctgc caagtttgcc	2220

# EP 1 310 571 A2

tcattttatca cgcagtagacag caccggccag gtcagcgtgg agatcgagtg ggaactgcag 2280  
 aaagaaaaca gcaaacgctg gaatccagag attcagtaca cctcaaatta tgccaagtct 2340  
 5 aataatgtgg aatttgctgt caacaacgaa ggggtttata ctgagcctcg ccccatgtggc 2400  
 acccggttacc tcacccgtaa cctgtaattg cctgttaatc aataaaccgg ttaattcgtt 2460  
 tcagttgaac tttggtcaag ggcaattc 2489  
 10  
 <210> 36  
 <211> 2495  
 <212> DNA  
 <213> new AAV serotype, clone 42.3b  
 15  
 <400> 36  
 gaattcgccc tttctacggc tgctgaact agaccaatga gaactttccc ttcaacgatt 60  
 gcgtcgacaa gatggtgatc tgggtggagg agggcaagat gacggccaag gtcgtggagt 120  
 20 ccgccaaggc cattcatcat ctgctggggc gggctccga gattgcttgc tcggcctgcg 180  
 atctggtcaa cgtggacctg gatgactgtg tttctgagca ataatgact taaaccaggt 240  
 atggctgccc atggttatct tccagattgg ctcgaggaca acctctctga gggcattcgc 300  
 25 gagtgggtggg acttgaaacc tggagccccg aaacccaaag ccaaccagca aaagcaggac 360  
 gacggccggg gtctggtgct tcctggctac aagtacctcg gaccctcaa cggactcgac 420  
 aagggagagc cgggtcaacga ggcagacgcc gcggccctcg agcacgacaa ggcctacgac 480  
 30 aagcagctcg agcaggggga caacccgtac ctcaagtaca accacgccga cgccgagttt 540  
 caggagcgtc ttcaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag 600  
 gccaagaagc gggttctcga acctctcggt ctggttgagg aaggcgctaa gacggctcct 660  
 35 ggaaagaaga gaccataga atccccgac tcctccacgg gcatcgga gaaaggccag 720  
 cagcccgcta aaaagaagct caactttggg cagactggcg actcagagtc agtgcccgac 780  
 cctcaaccaa tcggagaacc ccccgaggc ccctctggtc tgggatctgg tacaatggct 840  
 40 gcaggcgggtg gcgctccaat ggcagacaat aacgaaggcg ccgacggagt gggtaatgcc 900  
 tccggaaatt ggcattgca ttccacatgg ctgggagaca gagtcacac caccagcacc 960  
 cgcacctggg ccctgcccac ctacaacaac cacctctaca agcagatata aagtacagac 1020  
 45 ggggctacca acgacaacca cttcttcggc tacagcacc cctggggcta ttttgacttc 1080  
 aacagattcc actgccactt ctaccacgt gactggcagc gactcatcaa caacaactgg 1140  
 ggattccggc ccagaaagct gcggttcaag ttgttcaaca tccagggtcaa ggaggtcacg 1200  
 50 acgaacgacg gcgttacgac catcgcta atacctacca gcacgattca ggtcttctcg 1260  
 gactcggagt accaactgcc gtacgtctc ggctctgagc accagggctg cctccctccg 1320  
 ttccctgagg acgtgttcat gattcctcag tacggatata tgactctaaa caacggcagt 1380  
 55 cagtctgtgg gacgttcctc cttctactgc ctggagtact ttccttctca gatgctgaga 1440

# EP 1 310 571 A2

acgggcaata actttgaatt cagctacacc tttaggaag tgcctttcca cagcagctat 1500

gcgcacagcc agagcctgga cgggtgatg aatccctca tcgaccagta cctgtactac 1560

5 ctggcccga cccagagcac tacgggtcc acaaggagc tgcagttcca tcaggtggg 1620

cccaacacca tggccgagca atcaaagaac tggctgccc gacctgtta tcggcagcag 1680

agactgtcaa aaaacataga cagcaacaac accagtaact ttgcctggac cggggccact 1740

10 aaataccatc tgaatggtag aaattcatta accaaccgg gcgtagccat ggccaccaac 1800

aaggacgacg aggaccagtt ctttcccatc aacggagtgc tggtttttg caaacgggg 1860

gctgccaaca agacaacgct ggaaaacgtg ctaatgacca gcgaggagga gatcaaaacc 1920

15 accaatcccg tggctacaga acagtacggt gtggtctcca gcaacctgca atcgtctacg 1980

gccggacccc agacacagac tgtcaacagc cagggggctc tgcccggcat ggtctggcag 2040

aaccgggacg tgtacctgca ggggtccatc tgggccaaaa ttcttcacac ggacggcaac 2100

20 tttcaccggt ctcccctgat gggcggattt ggactcaaac accgcctcc tcaaattctc 2160

atcaaaaaca ccccggtacc tgctaatacct ccagaggtgt ttactcctgc caagtttgcc 2220

tcatttatca cgcagtacag caccggccag gtcagcgtg agatcgagt ggaactgcag 2280

25 aaagaaaaca gcaaacgctg gaatccagag attcagtaca cctcaaatta tgccaagtct 2340

aataatgtgg aatttgcgtg caacaacgaa ggggtttata ctgagcctcg cccattggc 2400

30 acccgttacc tcaccgtaa cctgtaattg cctgttaatc aataaaccgg ttaattcgtt 2460

tcagttgaac tttggtctct gcgaaggcg aattc 2495

<210> 37

<211> 3098

35 <212> DNA

<213> new AAV serotype, clone 42.11

<400> 37

gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60

40 gcgtcgacaa gatgggtgatc tgggtggagg agggcaagat gacggccaag gtcgtggagt 120

ccgccaggc cattctcggc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcttccg 180

cccagatcga tcccaccccc gtgatcgtca cttccaacac caacatgtgc gccgtgattg 240

45 acgggaacag caccaccttc gagcaccagc agccgttaca agaccggatg ttcaaatttg 300

aactcaccg ccgtctggag cagcactttg gcaaggtag aaagcaggaa gtcaaagagt 360

tcttcgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg 420

50 gtggagccaa caagagaccc gccccgatg acgcgataa aagcgagccc aagcggcct 480

gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agtccggtg gactttgccg 540

acaggtacca aaacaaatgt tctcgtcacg cgggcattgct tcagatgctg tttccctgca 600

55 agacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg accggagact 660

EP 1 310 571 A2

	gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc	720
	ggaaactctg tgccattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct	780
5	gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca	840
	ggtatggctg ccgatggtta tcttcagat tggctcgagg acaacctctc tgagggcatt	900
	cgcgagtggg gggacttgaa acctggagcc ccgaaacca aagccaacca gcaaaagcag	960
10	gacgacggcc ggggtctggt gcttctctgg tacaagtacc tcggaccctt caacggactc	1020
	gacaagggag agccggtcaa cgcggcggac gcagcggccc tcgagcacga caaggcctac	1080
	gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag	1140
15	tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggag agcagtcttc	1200
	caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct	1260
	cctggaaaga agagacccat agaatcccc gactcctcca cgggcatcgg caagaaaggc	1320
20	cagcagcccg ctaaaaagaa gctcaacttt gggcagactg gcgactcaga gtcagtgtcc	1380
	gacctcaac caatcggaga accccccgca ggccccctctg gtctgggatc tgggtacaatg	1440
	gctgcaggcg gtggcgctcc aatggcagac aataacgaag gcgccgacgg agtgggtaat	1500
25	gcctccggaa attggcattg cgattccaca tggctgggag acagagtcac caccaccagc	1560
	accgcacct gggccctgcc cacctacaac aaccacctct acaagcagat atcaagtcag	1620
	agcggggcta ccaacgacaa ccacttcttc ggctacagca cccctcggg ctattttgac	1680
30	ttcaacagat tccactgcc cttctcacca cgtgactggc agcgactcat caacaacaac	1740
	tggggattcc ggcccagaaa gctgcggttc aagtgtttca acatccaggt caaggaggtc	1800
	acgacgaacg acggcggtac gaccatcgct aataacctta ccagcacgat tcaggctcttc	1860
35	tcggactcgg agtaccact gccgtacgtc ctcggtctct cgcaccaggg ctgcctccct	1920
	ccgttccttg cggacgtgtt catgattcct cagtacggat atctgactct aaacaacggc	1980
	agtcagtctg tgggacgttc ctcttctac tgctggagt actttccttc tcagatgtg	2040
40	agaacgggca ataactttga attcagctac acctttgagg aagtgccttt ccacagcagc	2100
	tatgcgcaca gccagagcct ggaccggctg atgaatcccc tcacgacca gtacctgtac	2160
	tacctggccc ggaccagag cactacggg tccacaaggg agctgcagtt ccatcaggct	2220
45	gggcccaca ccatggccga gcaatcaaag aactggctgc ccggaccctg ttatcggcgg	2280
	cagagactgt caaaagacat agacagcaac aacaacagta actttgcctg gaccggggcc	2340
	actaaatacc atctgaatgg tagaaattca ttaaccaacc cgggcgtagc catggccacc	2400
50	aacaaggacg acgaggacca gttctttccc atcaacggag tgctgggttt tggcaaacg	2460
	ggggctgcc acaagacaac gctggaaaac gtgctaata ccagcgagga ggagatcaaa	2520
55	accaccaatc ccgtggctac agaagaatac ggtgtggtct ccagcaacct gcaatcgtct	2580

# EP 1 310 571 A2

acggccggac cccagacaca gactgtcaac agccaggggg ctctgcccgg catggtcttg 2640

cagaaccggg acgtgtacct gcaggggtccc atctgggcca aaattcctca cacggacggc 2700

5 aactttcacc cgtctcccct gatgggcgga tttggactca aacacccgcc tcctcaaatt 2760

ctcatcaaaa acaccccggt acctgctaatt cctccagagg tgtttactcc tgccaagttt 2820

gcctcattta tcacgcagta cagcaccggc caggtcagcg tggagatcga gtgggaactg 2880

10 cagaaagaga acagcaaacg ctggaatcca gagattcagt acacctcaaa ttatgccaaag 2940

tctaataatg tggaatttgc tgtcaacaac gaagggggtt atactgagcc tcgccccatt 3000

ggcacccggt acctcaccgg taacctgtaa ttacttggtta atcaataaac cggttgatgc 3060

15 gtttcagttg aactttgggc tctgcgaagg gcgaattc 3098

<210> 38

<211> 3276

20 <212> DNA

<213> new AAV serotype, clone 42.6a

<400> 38

gaattcgccc ttcgcagaga ccaaagttca actgaaacga attaacgggt ttattgatta 60

25 acaggcaatt acaggttacg ggtgaggtaa cgggtgccaa tggggcgagg ctcaagtataa 120

accccttcgt tgttgacagc aaattccaca ttattagact tggcataatt tgaggtgtac 180

tgaatctctg gattccagcg tttgctgttt tctttctgca gttccactc gatctccacg 240

30 ctgacctggc cgggtgctgta ctgcgtgata aatgaggcaa acttggcagg agtaaacacc 300

tctggaggat tagcaggtag cgggggtgtt ttgatgagaa tttgaggagg cgggtgtttg 360

agtccaaatc cgtccatcag gggagacggg tgaagttgc cgtccgtgtg aggaattttg 420

35 gccagatgg gaccctgcag gtacacgtcc cggttctgcc agaccatgcc gggcagagcc 480

ccctggctgt tgacagtctg tgtctggggc cgggccgtag acgattgcag gttgctggag 540

accacaccgt attcttctgt agccacggga ttggtgggtt tgatctctc ctgctggtc 600

40 attagcacgt tttccagcgt tgtctgtttg gcagccccg ttttgccaaa aaccagcact 660

cggttgatgg gaaagaactg gtcctcgtcg tccttggttg tggccatggc tacgccccgg 720

ttggttaatg aatttctacc attcacatgg tatttagtgg ccccggtcca ggcaaagtta 780

45 ctggtgttgt tgctgtctat gttttttgac agtctctgct gccgataaca ggggtccgggc 840

agccagttct ttgattgctc ggccatggtg ttgggccag cctgatggaa ctgcagctcc 900

cttgtggacc ccgtagtgct ctgggtccgg gccaggtagt acaggtactg gtcgatgagg 960

50 ggattcatca gccgggtccag gctctggcta tgcgcatagc tgctgtggaa aggcacttcc 1020

tcaaagggtg agctgaattc aaagttattg cccgttctca gcactctgaga aggaaagtac 1080

tccaggcagt agaaggagga acgtccaca gactgactgc cgttgtttag agtcagatat 1140

55 ccgtactgag gaatcatgaa cacgtccgca gggaacggag ggaggcagcc ctggtgcgca 1200

EP 1 310 571 A2

	gagccgagga cgtacggcag ttggtactcc gagtccgaga agacctgaat cgtgctggtg	1260
	aggttatttag cgatggctcg aacgccgtcg tccgtcgtga cctccttgac ctggatgttg	1320
5	aacaacttga accgcagctt tctgggccgg aatccccagt tgttgttgat gagtccgtgc	1380
	cagtcacgtg gtgagaagtg gcagtggaat ctgttaaagt caaaataccc ccagggggtg	1440
10	ctgtagccga agtaggtgtt gtggttggtg ctctctcccg atgtcccgtt ggagatttgc	1500
	ttgtagaggt ggttgttgta ggtggggagg gccaggttc ggtgctggt ggtgatgact	1560
	ctgtcgccca gccatgtgga atcgcaatgc caatttcctg aggaactacc cactccgtcg	1620
15	gcgccttcgt tattgtctgc cattggagcg ccaccgctg cagccattgt accagatccc	1680
	agaccagagg ggcctgcggg gggttctccg attggttgag ggtcgggcac tgactctgag	1740
	tcgccagtct gccc aaagt gagtctcttt ttcgcgggct gctggcctgt cttgccgatg	1800
20	cccgtagagg agtctggaga acgctgggt gatggctcta ccggtctctt cttccagga	1860
	gccgtcttag cgccttcctc aaccagaccg agaggttcga gaaccgctt cttggcctgg	1920
	aagactgctc gcccgagggt gccccaaaa gacgtatctt cttgaagacg ctccctgaaac	1980
25	tcggcgctcg cgtggttgta cttgaggtac gggttgtccc cctgctcgag ctgcttgctg	2040
	taggccttgt cgtgctcgag ggccgcggcg tctgcctcgt tgaccggctc tcccttgctg	2100
	agtccttgta agggctccgag gtactttag ccaggaagca ccagaccccg gccgtcgtcc	2160
30	tgcttttgct ggttggtctt gggtttcggg gctccaggtt tcaagtcca cactcgcga	2220
	atgccctcag agaggttgct ctcgagcaa tctggaagat aaccatcggc agccatacct	2280
	ggtttaagtc atttattgct cagaaacaca gtcattccagg tccacgttga ccagatcgca	2340
35	ggccgagcaa gcaatctcgg gagcccgccc cagcagatga tgaatggcac agagtttccg	2400
	atacgtcctc tttctgacga ccggttgaga ttctgacacg ccgggggaaac attctgaaca	2460
	gtctctggtc ccgtgcgtga agcaaatgtt gaaattctga ttcattctct cgcattgctt	2520
40	gcaggggaaac agcatctgaa gcatgccgc gtgacgagaa cacttgtttt ggtacctgtc	2580
	ggcaaagtcc accggagctc ctccgcgtc tgacgtcgat ggatgcaaaa tgtcgcaaaa	2640
	gcactcacgt gacagctaatacaggaccac tcccctatga cgtgatattac gtcagcgcta	2700
45	tgcccgctg acgagaacat ttgttttggt acctgtcggc aaagtccacc ggagctcctt	2760
	ccgcgtctga cgtcgatgga tccgcgactg aggggcaggc ccgcttgggc tcgcttttat	2820
	ccgcgtcatc gggggcggt ctcttggttg ctccaccctt tctgacgtag aactcatgcg	2880
50	ccacctcggc cagtgatcc tgcgccagc ggaagaactc tttgacttcc tgctttgtca	2940
	ccttgccaaa gtcatgctcc agacggcggg tgagttcaaa tttgaacatc cggtcctgca	3000
	acggctgctg gtgctcgaag gtggtgctgt tcccgctcaat cacggcgcac atgttggtgt	3060
55	tggaagtgac gatcacggg gtgggatcga tctgggcgga agacttgcac ttttggtcca	3120

# EP 1 310 571 A2

cgcgacacctt gctgccgccg agaatggcct tggcggactc cacgaccttg gccgtcatct 3180

tgccctcctc ccaccagatc accatcttgt cgacgcaatc gttgaaggga aagttctcat 3240

5 tgggccagtt gacgcagccg tagaaagggc gaattc 3276

<210> 39

<211> 3084

10 <212> DNA

<213> new AAV serotype, clone 43.1

<400> 39

gaattcgccc tttctacggc tgcatcaact ggaccaatga gaactttccc ttcaacgatt 60

15 gcgtcgacaa gatggtgatc tggatggagg agggcaagat gacggccaag gtcgtggagt 120

ccgccaaaggc cattctcggc ggcagcaagg tcgcgctgga ccaaaagtgc aagtcgtccg 180

cccagatcga cccaccccc gtgatcgtca cctccaacac caacatgtgc gccgtgattg 240

20 acgggaacag caccaccttc gagcaccagc agccgttgca ggaccggatg ttcaagtctg 300

aactcaccgc ccgtctggag cacgactttg gcaagggtgac caagcaggaa gtcaaagagt 360

tcttcgctg ggcgcaggat cacgtgaccg aggtggcgca tgagtcttac gtcagaaagg 420

25 gcggagccag caaaagaccc gccccgatg acgcggatat aagcgagccc aagcgggcct 480

gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540

acaggtagca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca 600

30 aaacgtgcga gaaaatgaat cagaatttca acatttgctt cacgcacggg gtcagagact 660

gctcagaatg tttccccggt gcatcagaat ctcaaccggt cgtcagaaaa aaaacgtatc 720

agaaactgtg tgccattcat catctgctgg ggcgggcacc cgagattgct tgctcggcct 780

35 gcgatctggt caacgtggac ctggacgact gtgtttctga gcaataaatg acttaaacca 840

ggtatggctg ccgatgggta tcttcagat tggcttgagg acaacctctc tgagggcatt 900

cgcgagtggg gggacctgaa acctggagcc ccgaaacca aagccaacca gcaaaagcag 960

40 gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc 1020

gacaaggggg agcccgtaa cgcggcggac gcagcggccc tcgagcacga caaggcctac 1080

gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag 1140

45 tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200

caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260

cctggaaaga agagaccggt agagccatca cctcagcgtt cccccgactc ctccacgggc 1320

50 atcggcaaga aaggccacca gcccgcgaga aagagactga actttgggca gactggcgac 1380

tcggagttag tccccgacct tcaaccaatc ggagaaccac cagcaggccc ctctggctctg 1440

ggatctggta caatggctgc aggcggtggc gctccaatgg cagacaataa cgaaggcgcc 1500

55 gacggagtgg gtagttcctc aggaaattgg cattgcgatt ccacatggct gggcgacaga 1560

# EP 1 310 571 A2

	gtcatcacca ccagcacccg aacctggggc ctgcccacct acaacaacca tctctacaag	1620
	caaattctcca acgggacatc gggaggaagc actaacgaca acacctactt tggctacagc	1680
5	accccctggg ggtattttga cttcaacaga ttccactgcc acttctcacc acgtgactgg	1740
	cagcgactca tcaacaataa ctggggattc cggcccaaga gactcaactt caagctcttc	1800
	aacatccagg tcaaggaggt cagcgagaat gaaggcacca agaccatcgc caataacctt	1860
10	accagcacga ttcagggtgtt tacggactcg gaataccagc tcccgtacgt ccccggtctt	1920
	gcgccaccagg gctgcctccc tccgttcccc gcggaagtct tcatgattcc tcagtacggg	1980
	tatctgaccc taaacaatgg cagtcaggct gtggggcgtt cctccttcta ctgcctggaa	2040
15	tacttccctt ctcaaagtct gaggacgggc aacaactttg aattcagcta caccttcgag	2100
	gacgtgcctt tccacagcag ctacgcgcac agccagagcc tggaccggct gatgaacctt	2160
	ctcatcgacc agtacctgta ttacttatcc agaactcagt ccacaggagg aactcaaggt	2220
20	actcagcaat tgttattttc tcaagccggg cccgcaaaca tgtcgggtca ggccaagaac	2280
	tggctacctg gaccgtgtta ccgtcagcaa cgagtttcca cgacactgtc gcaaaacaac	2340
	aacagcaatt ttgcttggac cgggtgccacc aagtatcacc tgaatggcag agactccctg	2400
25	gttaatcccg gcgttgccat ggctaccac aaggacgacg aggagcgctt cttcccgta	2460
	agcggagttc taatgtttgg caagcagggg gctggaaaag acaatgtgga ctacagcagc	2520
	gtgatgctca ccagcgaaga agaaattaaa actactaacc cagtgggtac agagcagtat	2580
30	ggtgtggtgg cagacaacct gcagcagacc aacggagctc ccattgtggg aactgtcaac	2640
	agccaggggg ccttacctgg tatggtctgg caaaaccggg acgtgtacct gcagggcccc	2700
35	atctgggcca aaattcctca cacggacggc aactttcacc cttcgccgct gatgggaggc	2760
	tttggtactga aacacccgcc tcctcagatc ctggtgaaaa acactcctgt tcctgcggat	2820
	cctccgacca ccttcagcca ggccaagctg gcttctttta tcacgcagta cagcaccgga	2880
40	caggtcagcg tggaaatcga atgggagctg cagaaagaaa acagcaagcg ctggaacca	2940
	gagattcagt atacttccaa ctactacaaa tctacaaatg tggactttgc tgtcaatact	3000
	gagggacttt attcagagcc tcgccccatt ggcactcgtt atctcacccg taatctgtaa	3060
45	ttgcttgtta atcaataaac cggt	3084
	<210> 40	
	<211> 2370	
	<212> DNA	
50	<213> new AAV serotype, clone 43.5	
	<400> 40	
	gaattcgccc tttctacggc tgcgtaact ggaccaatga gaactttccc ttcaacgatt	60
	gcgtcgacaa gatgggtgatc tgggtgggagg agggcaagat gacggccaag gtcgtggagt	120
55	ccgccaaggc cattctcggc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcgtccg	180

EP 1 310 571 A2

	cccagatcga cccaccccc gtgatcgta cctccaacac caacatgtgc gccgtgattg	240
	acgggaacag caccaccttc gagcaccagc agccgttgca ggaccggatg ttcaagttcg	300
5	aactcaccog ccgtctggag cagcactttg gcaaggtgac caagcaggaa gtcaaagagt	360
	tcttccgctg ggcgaggat cagtgaccg aggtggcgca tgagttctac gtcagaaagg	420
	gctggagccag caaaagaccc gcccccgatg acgcggtat aagcgagccc aagcgggcct	480
10	gcccctcagt cgcggtacca tcgacgtcag acgcggaagg agctccggtg gactttgccg	540
	acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagacgctg tttccctgca	600
	aaacgtgcga gagaatgaat cagaatttca acatttgctt cagcacggg gtcagagact	660
15	gctcagaatg tttccccggt gcatcagaat ctcaaccggt cgtcagaaaa aaaacgtatc	720
	agaaactgtg tgccattcat catctgctgg ggcgggcacc cgagattgct tgctcggcct	780
	gctgctggt caacgtggac ctggacgact gtgtttctga gcaataaatg acttaaacca	840
20	ggtatggctg ccgatggta tcttcagat tggcttgagg acaacctctc tgagggcatt	900
	cgcgagtggg gggacctgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag	960
	gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc	1020
25	gacaaggggg agcccgtaaa cgcgggcgac gcagcggccc tcgagcacga caaggcctac	1080
	gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag	1140
	tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcgggag agcagtcttc	1200
30	caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct	1260
	cctggaaaga agagaccggt agagccatca cctcagcggt cccccgactc ctccacgggc	1320
	atcggcaaga aaggccacca gcccgcgaga aagagactga actttgggca gactggcgac	1380
35	tcggagttag tccccgaccc tcaaccaatc ggagaaccac cagcaggccc ctctggtctg	1440
	ggatctggta caatggctgc aggcggtggc gctccaatgg cagacaataa cgaaggcgcc	1500
	gacggagtgg gtagttcttc aggaatttg cattgcgatt ccacatggct gggcgacaga	1560
40	gtcatcacca ccagcaccog aacctgggccc ctgcccacct acaacaacca tctctacaag	1620
	caaatctcca acgggacatc gggaggaagc actaacgaca acacctactt tggctacagc	1680
	acccctcggg ggtattttga cttcaacaga ttccactgcc acttctcacc acgtgactgg	1740
45	cagcgactca tcaacaataa ctggggattc cggcccaaga gactcaactt caagctcttc	1800
	aacatccagg tcaaggaggt cagcagaat gaaggcacca agaccatcgc caataacctt	1860
	accagcacga ttcaggtggt tacggactcg gaataccagc tcccgtacgt cctcggctct	1920
	gcgcaccagg gctgcctccc tccgttcccg gcggacgtct tcatgattcc tcagtaaggg	1980
	tatctgaccc taaacaatgg cagtcaggct gtgggccggt cctccttcta ctgctggaa	2040
55	tacttcctt ctcaaatgct gaggacgggc aacaactttg aattcagcta caccttcgag	2100

# EP 1 310 571 A2

gacgtgcctt tccacagcag ctacgcgcac agccagagcc tggaccggct gatgaaccct 2160  
ctcatcgacc agtacctgta ttacttatcc agaactcagt ccacaggagg aactcaaggt 2220  
5 actcagcaat tggtattttc tcaagccggg cccgcaaaca tgtyggctca ggccaagaac 2280  
tggctacctg gaccgtgta cgcgcagcaa cgagtttcca cgacactgtc gcaaaacaac 2340  
10 aacagcaatt ttgctggacc ggtgccacca 2370

<210> 41  
<211> 3123  
<212> DNA  
<213> new AAV serotype, clone 43.12

15 <400> 41  
gaattcgccc ttggctgcgt caactggacc aatgagaact ttcccttcaa cgattgcgtc 60  
gacaagatgg tgatctggtg ggaggagggc aagatgacgg ccaaggctcg ggagtccgcc 120  
20 aaggccattc tcggcggcag caaggtgcgc gtggaccaa agtgcaagtc gtccgccag 180  
atcgacccca ccccggtgat cgtcacctcc aacaccaaca tgtgcgccgt gattgacggg 240  
aacagcacca ccttcgagca ccagcagccg ttgcaggacc ggatgttcaa gttcgaactc 300  
25 acccgccgtc tggagcacga ctttggaag gtgaccaagc aggaagtcaa agagttcttc 360  
cgctgggcgc aggatcacgt gaccgaggtg gcgcattgagt tctacgtcag aaagggcgga 420  
gccagcaaaa gaccgcgccc cgatgacgcg gatataagcg agcccaagcg ggccctgccc 480  
30 tcagtcgcgg atccatcgac gtcagacgcg gaaggagctc cggtggaatt tgccgacagg 540  
taccaaaaca aatgttctcg tcacgcgggc atgctccaga tgctgtttcc ctgcaaaacg 600  
tgcgagagaa tgaatcagaa tttcaacatt tgcttcacgc acgggggtcag agactgctca 660  
35 gaatgtttcc ccggtgcac agaattctcaa ccggtcgtca gaaaaaaac gtatcagaaa 720  
ctgtgtgcca ttcacatct gctggggcgg gcacccgaga ttgcttgctc ggccctgcgat 780  
ctggtcaacg tggacctgga cgactgtgtt tctgagcaat aatgactta aaccaggat 840  
40 ggctgccgat ggttatcttc cagattggct tgaggacaac ctctctgagg gcattcgga 900  
gtggtgggac ctgaaacctg gagccccgaa acccaaagcc aaccagcaaa agcaggacga 960  
cgccgggggt ctggtgcttc ctggctacaa gtacctcgga cccttcaacg gactcgacaa 1020  
45 gggggagccc gtcaacgcgg cggacgcagc ggccctcgag caccgacaag cctacgacca 1080  
gcagctcaaa gcgggtgaca atccgtacct gcggtataac caccgcgacg ccgagtttca 1140  
ggagcgtctg caagaagata cgtcttttgg gggcaacctc gggcgagcag tcttcaggc 1200  
50 caagaagcgg gttctcgaac ctctcggctt ggttgaggaa ggcgctaaga cggctcctgg 1260  
aaagaagaga ccggtagagc catcacctca gcgttcccc gactcctcca cgggcatcgg 1320  
caagaaaggc caccagcccg cgagaaagag actgaacttt gggcagactg gcgactcgga 1380  
55 gtcagtcccc gaccctcaac caatcggaga accaccagca ggccctctg gtctgggatc 1440

# EP 1 310 571 A2

	tggtacaatg gctgcaggcg gtggcgctcc aatggcagac aataacgaag gcgccgacgg	1500
5	agtgggtagt tccctcaggaa attggcattg cgattccaca tggctgggcg acagagtcac	1560
	caccaccagc acccgaacct ggccctgcc cacctacaac aaccatctct acaagcaaat	1620
	ctccaacggg acatcgggag gaagcactaa cgacaacacc tactttggct acagcacccc	1680
10	ctgggggtat tttgacttca acagattcca ctgccacttc tcaccacgtg actggcagcg	1740
	actcatcaac aataactggg gattccggcc caagagactc aacttcaagc tcttcaacat	1800
	ccagggtcaag gaggtcacgc agaatgaagg caccaagacc atcgccaata accttaccag	1860
15	cacgattcag gtgtttacgg actcgggaata ccagctcccg tacgtcctcg gctctgcgca	1920
	ccagggtctgc ctccctccgt tccggcgga cgtcttcacg attcctcagt acgggtatct	1980
	gaccctaaac aatggcagtc aggtctggg ccgttcctcc ttctactgcc tggaataactt	2040
20	cccttctcaa atgctgagga cgggcaacaa ctttgaattc agctacacct tcgaggacgt	2100
	gcctttccac agcagctacg cgcacagcca gagcctggac cggctgatga accctctcat	2160
	cgaccagtac ctgtattact tatccagaac tcagtccaca ggaggaactc aaggtactca	2220
25	gcaattgtta ttttctcaag ccgggcccgc aaacatgtcg gctcaggcca agaactggct	2280
	acctggaccg tgttaccgtc agcaacgagt ttccacgaca ctgtcgcaa acaacaacag	2340
	caattttgct tggaccggtg ccaccaagta tcacctgaat ggcagagact ccctgggtta	2400
30	tccggcggt gccatggcta cccacaagga cgacgaggag cgcttcttcc cgtcaagcgg	2460
	agttctaatg tttggcaagc agggggctgg aaaagacaat gtggactaca gcagcgtgat	2520
	gctcaccagc gaagaagaaa ttaaaactac taaccacgtg gctacagagc agtatgggtg	2580
35	ggtggcagac aacctgcagc agaccaacgg agctccatt gtgggaactg tcaacagcca	2640
	gggggcctta cctggtatgg tctggcaaaa ccgggacgtg tacctgcagg gcccacatctg	2700
	ggccaaaatt cctcacacgg acggcaactt tcactcttcg ccgctgatgg gaggttttgg	2760
40	actgaaacac ccgctcctc agatcctggg gaaaaacact cctgttcctg cggatcctcc	2820
	gaccaccttc agccaggcca agctggcttc ttttatcacg cagtacagca ccggacaggt	2880
	cagcgtggaa atcgaatggg agctgcagaa agaaaacagc aagcgtgga acccagagat	2940
45	tcagtatact tccaactact acaaactctac aaatgtggac tttgctgtca atactgaggg	3000
	tacttattca gagcctcgcc ccattggcac tcgttatctc acccgtaatc tgtaattgct	3060
	tgттаатcaa taaaccggtt aattcgtttc agttgaactt tggctctctgc gaagggcgaa	3120
50	ttc	3123

<210> 42  
 <211> 3122  
 <212> DNA  
 <213> new AAV serotype, clone 43.20

EP 1 310 571 A2

<400> 42  
gaattcggccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60  
5 gcgctgcacaa gatggtgata tgggtggagg agggcaagat gacggccaag gtcgtggagt 120  
ccgccaaggc cattctcggc ggcagcaagg tgcgtgtgga ccaaaagtgc aagtcttccg 180  
cccagatcga tcccccccc gtgatcgtca cctccaacac caacatgtgc gccgtgattg 240  
10 acgggaacag cgccaccttc gagcaccagc agccgttgca ggaccggatg ttcaaatttg 300  
aactcaccgc ccgtctggag catgactttg gcaaggtgac gaagcaggaa gtcaaagagt 360  
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagtccac gtcagaaagg 420  
15 gtggagccaa caagagaccc gccccgatg acgcggatat aagcgagccc aagcgggcct 480  
gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540  
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca 600  
20 agacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg accagagact 660  
gttcagaatg tttccccggc gtgtcagaat ctcaaccggg cgtcagaaag aggacgtatc 720  
ggaaactctg tgcgattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct 780  
25 gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaacca 840  
ggtatggctg ccgatggta tcttcagat tggctcgagg acaacctctc tgagggcatt 900  
cgcgagtggg gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag 960  
30 gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc 1020  
gacaaggggg agcccgtaaa cgcggcggac gcagcggccc tcgagcacga caaagcctac 1080  
gaccagcagc tcaaagcggg tgacaatccg tacctgcggg ataatacagc cgacgccgag 1140  
35 tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcggggc agcagtcttc 1200  
caggccaaga agcgggttct cgaacctctc ggtctggtt aggaaggcgc taagacggct 1260  
cctggaaaga agagactggt agagcagtcg ccacaagagc cagactcctc ctcgggcatc 1320  
40 ggcaagacag gccagcagcc cgctaaaaag agactcaatt ttggtcagac tggcgactca 1380  
gagtcagtcc ccgaccaca acctctcgga gaacctccag cagccccctc aggtctggga 1440  
cctaatacaa tggcttcagg cgggtggcgt ccaatggcag acaataacga aggcgccgac 1500  
45 ggagtgggta attcctcggg aaattggcat tgcgattcca catggctggg ggacagagtc 1560  
atcaccacca gcacccgaac ctgggccctg cccacctaca acaaccacct ctacaagcaa 1620  
atctccaacg gcacctcggg aggaagcacc aacgacaaca cctatttttg ctacagcacc 1680  
50 ccctgggggt attttgactt caacagattc cactgtcact tttcaccagc tgactggcaa 1740  
cgactcatca acaacaattg gggattccgg cccaaaagac tcaacttcaa gctgttcaac 1800  
atccaggtca aggaagtcac gacgaacgaa ggcaccaaga ccatcgccaa taatctcacc 1860  
55 agcacctgtc aggtctttac ggactcggag taccagttac cgtacgtgct aggatccgct 1920

# EP 1 310 571 A2

	caccagggat gtctgcctcc gttcccggcg gacgtcttca cggttcctca gtacggctat	1980
5	ttaacttttaa acaatggaag ccaagccctg ggacgttcct ccttctactg tctggagtat	2040
	ttcccatcgc agatgctgag aaccggcaac aactttcagt tcagctacac cttcgaggac	2100
	gtgcctttcc acagcagcta cgcgcacagc cagagcctgg acaggtgat gaatcccctc	2160
10	atcgaccagt acctgtacta cctggtcaga acgcaaacga ctggaactgg agggacgcag	2220
	actctggcat tcagccaagc gggtcctagc tcaatggcca accaggctag aaattgggtg	2280
	cccggacctt gctaccggca gcagcgcgtc tccacgacaa ccaaccagaa caacaacagc	2340
15	aactttgcct ggacgggagc tgccaagttt aagctgaacg gccgagactc tctaataaat	2400
	ccgggctgtg caatggcttc ccacaaggat gacgacgacc gcttcttccc ttcgagcggg	2460
	gtcctgattt ttggcaagca aggagccggg aacgatggag tggattacag ccaagtgtgt	2520
20	attacagatg aggaagaaat caaggctacc aaccccgtgg ccacagaaga atatggagca	2580
	gtggccatca acaaccaggc cgccaatacg caggcgcaga ccggactcgt gcacaaccag	2640
	ggggtgattc ccggcatggt gtggcagaat agagacgtgt acctgcaggg tcccatctgg	2700
25	gccaaaattc ctcacacgga cggcaacttt caccctcttc ccctgatggg cggctttgga	2760
	ctgaagcacc cgctcctca aattctcatc aagaacacac cggttccagc ggacccgccg	2820
	cttaccttca accaggccaa gctgaactct ttcacacgc agtacagcac cggacaggtc	2880
30	agcgtggaaa tcgagtggga gctgcagaaa gaaaacagca aacgctggaa tccagagatt	2940
	caatacactt ccaactacta caaatctaca aatgtggact ttgctgtcaa cacggaagga	3000
	gtttatagcg agcctcgccc cattggcacc cgttacctca cccgcaacct gtaattacat	3060
35	gttaatcaat aaaccggtta attcgtttca gttgaacttt ggtctctgcg aagggcgaat	3120
	tc	3122
40	<210> 43 <211> 3117 <212> DNA <213> new AAV serotype, clone 43.21	
45	<400> 43 gaattcgccc ttggctgcgt caactggacc aatgagaact ttcccttcaa cgattgcgtc	60
	gacaagatgg tgatctggtg ggaggagggc aagatgacgg ccaaggctcg ggagtccgcc	120
	aaggccattc tcggcggcag caagggtcgt gtggaccaa agtgcaagtc ttccgccag	180
50	atcgatccca cccccgtgat cgtcacctcc aacaccaaca tgtgcgccgt gattgacggg	240
	aacagcacca ccttcgagca ccagcagccg ttgcaggacc ggatgttcaa atttgaactc	300
	acccgccgtc tggagcatga ctttggcaag gtgacgaagc aggaagtcaa agagttcttc	360
55	cgctgggctc aggatcacgt gaccgaggtg gcgcatgagt tccacgtcag aaagggtgga	420

EP 1 310 571 A2

	gccaacaaga gacccgcccc cgatgacgcg gatataagcg agcccaagcg ggccctgcccc	480
	tcagtgcgcg atccatcgac gtcagacgcg gaaggagctc cgggtggactt tgccgacagg	540
5	tacaaaaaca aatgttctcg tcacgcgggc atgcttcaga tgctgtttcc ctgcaagaca	600
	tgcgagagaa tgaatcagaa tttcaacatt tgcttcacgc acgggaccag agactgttca	660
	gaatgtttcc ccggcgtgtc agaattctca ccggtcgtca gaaagaggac gtatcggaaa	720
10	ctctgtgcga ttcatcatct gctggggcgg gctcccgaga ttgcttgctc ggccctgcat	780
	ctggtcaacg tggacctgga tgactgtgtt tctgagcaat aaatgactta aaccaggat	840
	ggctgccgat ggttatcttc cagattggct cgaggacaac ctctctgagg gcattcgcga	900
15	gtggtgggac ttgaaacctg gagccccgaa acccaaagcc aaccagcaa agcaggacga	960
	cggccgggggt ctggtgcttc ctggctacaa gtacctcgga cccttcaacg gactcgacaa	1020
	gggggagccc gtcaacgcgg cggacgcagc ggccctcgag cagacaaaag cctacgacca	1080
20	gcagctcaaa gcgggtgaca atccgtacct gcggtataat cagccgacg ccgagtttca	1140
	ggagcgtctg caagaagata cgtcttttgg gggcaacctc gggcgagcag tcttcaggc	1200
	caagaagcgg gttctcgaac ctctcgggtc ggttgaggaa ggcgctaaga cggctcctgg	1260
25	aaagaagaga ccggtagagc agtcgccaca agagccagac tcctcctcgg gcatcggcaa	1320
	gacaggccag cagcccgtc aaaagagact caattttggt cagactggcg actcagagtc	1380
	agtcctcgac ccacaacctc tcggagaacc tcagcagcc ccctcaggtc tgggacctaa	1440
30	tacaatggct tcaggcgggtg gcgtctcaat ggagacaaat aacgaaggcg ccgacggagt	1500
	gggtaattcc tcgggaaatt ggcatcgga ttccacatgg ctgggggaca gagtcatcac	1560
	caccagcacc cgaacctggg ccctgcccac ctacaacaac cacctctaca agcaaattctc	1620
35	caacggcacc tcgggaggaa gcaccaacga caacacctat tttggctaca gcaccccctg	1680
	ggggatattt gacttcaaca gattccactg tcaattttca ccacgtgact ggcaacgact	1740
	catcaacaac aattggggat tccggcccaa aagactcaac ttcaagctgt tcaacatcca	1800
40	ggtcaaggaa gtcacgacga acgaaggcac caagaccatc gccaataatc tcaccagcac	1860
	cgtgcgggtc tttacggact cggagtacca gttaccgtac gtgctaggat ccgtcacca	1920
	gggatgtctg cctccgttcc cggcggacgt ctcatgggt cctcagtacg gctatttaac	1980
45	tttaaacaaat ggaagccaag ccctgggacg ttctctcttc tactgtctgg agtatttccc	2040
	atcgagatg ctgagaaccg gcaacaactt tcagttcagc tacaccttcg aggacgtgcc	2100
	tttccacagc agctacgcgc acagccagag cctggacagg ctgatgaatc ccctcatcga	2160
50	ccagtacctg tactacctgg tcagaacgca aacgactgga actggaggga cgcagactct	2220
	ggcattcagc caagcgggtc ctagctcaat ggccaaccag gctagaaatt ggggtgcccgg	2280
55	accttgctac cggcagcagc gcgtctccac gacaaccaac cagagcaaca acagcaactt	2340

# EP 1 310 571 A2

	tgcttgacg ggagctgcca agtttaagct gaacggccga gactctctaa tgaatccggg	2400
	cgtggcaatg gcttcccaca aggatgacga cgaccgcttc ttcccttcga gcgggggtcct	2460
5	gatttttggc aagcaaggag ccgggaacga tggagtggat tacagccaag tgctgattac	2520
	agatgaggaa gaaatcaagg ctaccaaccc cgtggccaca gaagaatatg gagcagtggc	2580
10	catcaacaac caggccgcca atacgcaggc gcagaccgga ctcgtagcaca accaggggggt	2640
	gattccccgc atggtgtggc agaatagaga cgtgtacctg caggggtccca tctgggccaa	2700
	aattcctcac acggacggca actttcaccc gtctcccctg atggggcggct ttggactgaa	2760
15	gcacccgcct cctcaaattc tcatcaagaa cacaccggtt ccagcggacc cgccgcttac	2820
	cttcaaccag gccaaactga actctttcat cagcgagtac agcaccggac aggtcagcgt	2880
	ggaaatcgag tgggagctgc agaaagaaaa cagcaaacgc tggaatccag agattcaata	2940
20	cacttccaac tactacaaat ctacaaatgt ggacttttgc gtcaacacgg aaggagttta	3000
	tagcgagcct cgccccattg gcacccgtta cctcacccgc aacctgtaat tacatgttaa	3060
	tcaataaacc ggttaattcg tttcagttga actttggctc ctgcgaaggg cgaattc	3117
25	<210> 44	
	<211> 3121	
	<212> DNA	
	<213> new AAV serotype, clone 43.23	
30	<400> 44	
	gaattcgccc ttctacggct gcgtcaactg gaccaatgag aactttccct tcaacgattg	60
	cgtagacaag atggtgatct ggtgggagga gggcaagatg acggccaagg tcgtggagtc	120
	cgccaaggcc attctcggcg gcagcaagggt gcgtgtggac caaaagtga agtcttccgc	180
35	ccagatcgat cccacccccg tgatcgtcac ctccaacacc aacatgtgcg ccgtgattga	240
	cggaacagc accaccttcg agcaccagca gccgttgacg gaccggatgt tcaaatttga	300
	actcacccgc cgtctggagc atgactttgg caaggtgacg aagcaggaag tcaaagagtt	360
40	cttccgctgg gcgcaggatc acgtgaccga ggtggcgcac gaggttccac tcagaaaggg	420
	tggcgccaac aagagacccg ccccgatga cgcgatata agcgagcca agcgggcctg	480
	cccctcagtc gcggatccat cgacgtcaga cgcggaagga gctccggtgg actttgccga	540
45	caggtaccaa aacaaatgtt ctcgtcacgc gggcatgctt cagatgctgt ttccctgcaa	600
	gacatgagag agaataatc agaatttcaa catttgcttc acgcacggga ccagagactg	660
	ttcagaatgt ttccccggcg tgcagaatc tcaaccggtc gtcagaaaga ggacgtatcg	720
50	gaaactctgt gcgattcatc atctgctggg gcgggctccc gagattgctt gctcggcctg	780
	cgatctggtc aacgtggacc tggatgactg tgtttctgag caataaatga cttaaaccag	840
	gtatggctgc cgatggttat cttccagatt ggctcgagga caacctctct gagggcatte	900
55	gcgagtgggtg ggacttgaac cctggagccc cgaaacccaa agccaaccag caaaagcagg	960

# EP 1 310 571 A2

	acgacggccg gggctctgggtg cttcctgggt acaagtacct cggacccttc aacggactcg	1020
	acaaggggga gcccgtcaac gcggcggacg cagcggccct cgagcacgac aaagcctacg	1080
5	accagcagct caaagcgggt gacaatccgt acctgcggta taatcacgcc gacgccgagt	1140
	ttcaggagcg tctgcaagaa gatacgtcct ttgggggcaa cctcgggcga gcagtcttcc	1200
	aggccaagaa gcgggttctc gaacctctcg gtctggttga ggaaggcgct aagacggctc	1260
10	ctggaaagaa gagaccggtg gagcagtcgc cacaagagcc agactcctcc tcgggcatcg	1320
	gcaagacagg ccagcagccc gctaaaaaga gactcaattt tggtcagact ggcgactcag	1380
	agtcagtccc cgaccacaaa cctctcggag aacctccagc agccccctca ggtctgggac	1440
15	ctaatacaat ggcttcaggc ggtggcgctc caatggcaga caataacgaa ggcgccgacg	1500
	gagtgggtaa ttcctcggga aattggcatt gcgattccac atggctgggg gacagagtca	1560
	tcaccaccag caccgaacc tgggccttgc ccacctaca caaccacctc tacaagcaaa	1620
20	tctccaacgg cacctcggga ggaagcacca acgacaacac ctattttggc tacagcacc	1680
	cctgggggta ttttgacttc aacagattcc actgtcactt ttcaccacgt gactggcaac	1740
	gactcatcaa caacaattgg ggattccggc ccaaaagact caacttcaag ctgttcaaca	1800
25	tccagggtcaa ggaagtcacg acgaacgaag gcaccaagac catcgccaat aatctcacca	1860
	gcaccgtgca ggtctttacg gacttggagt accagttacc gtacgtgcta ggatccgctc	1920
	accagggatg tctgcctccg ttcccggcgg acgtcttcat ggttcctcag tacggctatt	1980
30	taactttaaa caatggaagc caagccctgg gacgttcctc cttctactgt ctggagtatt	2040
	tcccatcgca gatgccgaga accggcaaca actttcagtt cagctacacc ttcgaggacg	2100
	tgcctttcca cagcagctac gcgcacagcc agagcctgga caggctgatg aatccccctca	2160
35	tcgaccagta cctgtactac ctggtcagaa cgcaaacgac tggaaactgga gggacgcaga	2220
	ctctggcatt cagccaagcg ggtcctagct caatggccaa ccaggctaga aattgggtgc	2280
	ccggaccttg ctaccggcag cagcgcgtct ccacgacaac caaccagaac aacaacagca	2340
40	actttgcctg gacgggagct gccaaagtta agctgaacgg ccgagactct ctaatgaatc	2400
	cgggcgtggc aatggcttcc cacaaggatg acgacgaccg cttcttccct tcgagcgggg	2460
	tcttgatttt tggcaagcaa ggagccggga acgatggagt ggattacagc caagtgtctga	2520
45	ttacagatga ggaagaaatc aaggctacca acccgtggc cacagaagaa tatggagcag	2580
	tggccatcaa caaccaggcc gccaatagc aggcgcagac cggactcgtg cacaaccagg	2640
	gggtgattcc cggcatgggtg tggcagaata gagacgtgta cctgcagggt cccatctggg	2700
50	ccaaaattcc tcacacggac ggcaactttc acccgtctcc cctgatgggc ggctttggac	2760
	tgaagcacc gcctcctcaa attctcatca agaacacacc ggttccagcg gaccgccgc	2820
55	ttaccttcaa ccaggccaag ctgaactctt tcatcacgca gtacagcacc ggacagggtca	2880

EP 1 310 571 A2

gcgtggaaat cgagtgggag ctgcagaaaag aaaacagcaa acgctggaat ccagagattc 2940

aatacacttc caactactac aaatctacaa atgtggactt tgctgtcaac acggaaggag 3000

5 tttatagcga gcctcgcccc attggcaccc gttacctcac ccgcaacctg taattacatg 3060

ttaatcaata aaccggttaa ttcgtttcag ttgaactttg gtctctgcga agggcgaatt 3120

c 3121

10 <210> 45  
<211> 3122  
<212> DNA  
<213> new AAV serotype, clone 43.25

15 <400> 45  
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60

gcgtcgacaa gatggtgatc tgggtggagg agggcaagat gacggccaag gtcgtggagt 120

ccgccaaggc cattctcggc ggcagcaagg tgcgtgtgga ccaaaagtgc aagtcttccg 180

20 cccagatcga tcccaccccc gtgatcgtca cctccaacac caacatgtgc gccgtgattg 240

acgggaacag caccaccttc gagcaccagc agccgttgca ggaccggatg ttcaaatttg 300

aactcaccgc ccgtctggag catgactttg gcaagggtgac gaagcaggaa gtcaaagggg 360

25 tcttcgcgtg ggcgaggat cactgtaccg aggtggcgca tgagttccac gtgcgagccc 420

aagcgggcct gccctcagt cgcggatcca tcgacgtcag accagaaagg gtggagccaa 480

caagagaccc gccccgatg acgcggatat aagcgaagg agctccggtg gactttgccg 540

30 acaggtacca aaacaaatgt tctcgtcagc cgggcatgct tcagatgctg tttccctgca 600

agacatgcga gagaatgaat cagaatttca acatttgctt cagcacggg accagagact 660

gttcagaatg tttccccggc gtgtcagaat ctcaaccggg cgtcagaaag aggacgtatc 720

35 ggaaactctg tgcgattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct 780

gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca 840

ggatattgctg ccgatgggta tcttcagat tggctcgagg acaacctctc tgagggcatt 900

40 cgcgagtggg gggacttgaa acctggagcc ccgaaacca aagccaacca gcaaaagcag 960

gacgacggcc ggggtctggt gcttccctggc tacaagtacc tcggaccctt caacggactc 1020

gacaaggggg agcccgtaaa cgcggcggac gcagcggccc tcgagcacga caaagcctac 1080

gaccagcagc tcaaagcggg tgacaatccg tacctgcggg ataatacacgc cgacgccgag 1140

tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcgggag agcagtcttc 1200

50 caggccaaga agcgggttct cgaacctctc ggtctgggtg aggaaggcgc taagacggct 1260

cctggaaaga agagaccggg agagcagtcg ccacaagagc cagactcctc ctcgggcatc 1320

ggcaagacag gccagcagcc cgctaaaaag agactcaatt ttggtcagac tggcgactca 1380

55 gagtcagtcc ccgaccaca acctctcgga gaacctccag cagccccctc aggtctggga 1440

# EP 1 310 571 A2

cctaatacaa tggcttcagg cgggtggcgct ccaatggcag acaataacga aggcgccgac 1500  
 ggagtgggta attcctcggg aaattggcat tgcgattcca catggctggg ggacagagtc 1560  
 5 atcaccacca gcacccgaac ctgggccctg cccacctaca acaaccacct ctacaagcaa 1620  
 atctccaacg gcacctcggg aggaagcacc aacgacaaca cctatttttg ctacagcacc 1680  
 ccctgggggt attttgactt caacagattc cactgtcact tttcaccacg tgactggcaa 1740  
 10 cgactcatca acaacaattg gggattccgg cccaaaagac tcaacttcaa gctgttcaac 1800  
 atccagggtca aggaagtcac gacgaacgaa ggcaccaaga ccatcgccaa taatctcacc 1860  
 agcacccgtgc aggtctttac ggactcggag taccagttac cgtacgtgct aggatccgct 1920  
 15 caccagggat gtctgcctcc gttcccgcg gacgtcttca tggttcctca gtacggctat 1980  
 ttaactttta acaatggaag ccaagccctg ggacgttctt ccttctactg tctggagtat 2040  
 ttcccatcgc agatgctgag aaccggcaac aactttcagt tcagctacac cttcgaggac 2100  
 20 gtgcctttcc acagcagcta cgcgcacagc cagagcctgg acaggctgat gaatcccctc 2160  
 atcgaccagt acctgtacta cctggtcaga acgcaaacga ctggaactgg agggacgcag 2220  
 actctggcat tcagccaagc gggtcctagc tcaatggcca accaggctag aaattgggtg 2280  
 25 ccgggacctt gctaccggca gcagcgcgtc tccacgacaa ccaaccagaa caacaacagc 2340  
 aactttgcct ggacgggagc tgccaagttt aagctgaacg gccgagactc tctaataaat 2400  
 ccggggcgtg caatggcttc ccacaaggat gacgacgacc gcttcttccc ttcgagcggg 2460  
 30 gtcttgattt ttggcaagca aggagccggg aacgatggag tggattacag ccaagtgtg 2520  
 attacagatg aggaagaaat caaggctacc aaccccgctg ccacagaaga atatggagca 2580  
 gtggccatca acaaccaggc cgccaatacg caggcgcaga ccggactcgt gcacaaccag 2640  
 35 ggggtgattc ccggcatggt gtggcagaat agagacgtgt acctgcaggg tcccatctgg 2700  
 gccaaaattc ctcacacgga cggcaacttt caccgcgtct ccctgatggg cggctttgga 2760  
 ctgaagcacc cgcctcctca aattctcatt aagaacacac cggttccagc ggacccgccg 2820  
 40 cttaccttca accaggccaa gctgaactct ttcacacgc agtacagcac cggacaggtc 2880  
 agcgtggaat tcgagtggga gctgcagaaa gaaaacagca aacgctggaa tccagagatt 2940  
 caatacactt ccaactacta caaatctaca aatgtggact ttgctgtcaa cacggagggg 3000  
 gtttatagcg agcctcgccc cattggcacc cgttacctca cccgcaacct gtaattacat 3060  
 gttaataaat aaaccggtta attcgtttca gttgaacttt ggtctctgcg aagggcgaat 3120  
 50 tc 3122

<210> 46  
 <211> 3128  
 <212> DNA  
 <213> new AAV serotype, clone 44.1

EP 1 310 571 A2

<400> 46  
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60  
5 gcgtcgacaa gatgttgatc tgggtggagg agggcaagat gacggccaag gtcgtggagt 120  
ccgccaaggc cattctcggc ggcagcaaag tgcgcgtgga ccaaaagtgc aagccgtccg 180  
cccagatcga cccaccccc gtgatcgtca cctccaacac caacatgtgc gccgtgattg 240  
10 acgggaacag caccaccttc gagcaccagc agccgttgcg ggaccggatg ttcaagtttg 300  
aactcaccgc ccgtctggag cactgactttg gcaagggtgac aaagcaggaa gtcagagagt 360  
tcttcgcgtg ggcgcaggat cactgtaccg aggtggcgca cgagttctac gtcagaaagg 420  
15 gtggagccaa caagagaccc gccccgatg acgcggataa aagcgagccc aagcgggcct 480  
gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agtccgggtg gactttgccg 540  
acaggtagca aaacaaatgt tctcgtcacg cgggcgtgct tcagatgctg tttccctgca 600  
20 aaacatgcga gagaatgaat cagaatttca acatttgctt cagcacggg accagagact 660  
gttcagaatg tttccccggc gtgtcagaat ctcaaccggg cgtcagaaaa aagacgtatc 720  
ggaaactctg tgcgattcat catctgctgg ggcgggcacc cgagattgct tgctcggcct 780  
25 gcgatctggt caacgtggac ctagatgact gtgtttctga gcaataaatg acttaacca 840  
ggtatggctg ccgatgggta tcttcagat tggctcgagg acaacctctc tgagggcatt 900  
cgcgagtggg gggacttgaa acctggagcc ccgaaacca aagccaacca gcaaaagcag 960  
30 gacgacggcc ggggtctggt gcttctctggc tacaagtacc tcggaccctt caacggactc 1020  
gacaaggggg agcccgtaaa cgcggcggac gcagcggccc tcgagcacga caaggcctac 1080  
gaccagcagc tcaaagcggg tgacaatccg tacctgcggg ataaccacgc cgacgccgag 1140  
35 tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcgggag agcagtcctc 1200  
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260  
cctggaaaga agagaccggg agagccatca cccagcgtt ctccagactc ctctacgggc 1320  
40 atcggcaaga aaggccagca gcccgcgaaa aagagactca actttgggca gactggcgac 1380  
tcagagtcag tgcccgacct tcaaccaatc ggagaacccc ccgagggccc ctctggtctg 1440  
ggatctggtg caatggctgc aggcggtggc gctccaatgg cagacaataa cgaaggcgcc 1500  
45 gacggagtgg gtagttcttc aggaatttgg cattgcgatt ccacatggct gggcgacaga 1560  
gtcatcacca ccagcaccgc aacctgggccc ctccccacct acaacaacca cctctacaag 1620  
caaatctcca acgggacttc gggaggaagc accaacgaca acacctactt cggctacagc 1680  
50 acccctggg ggtattttga ctttaacaga ttccactgcc acttctcacc acgtgactgg 1740  
cagcgactca tcaacaacaa ctggggattc cggcccaaga gactcaactt caagctcttc 1800  
aacatccag tcaaggaggt cactgcagaat gaaggcacca agaccatcgc caataacctt 1860  
55 accagcacga ttcaggtctt tacggactcg gaataccagc tcccgtacgt cctcggctct 1920

# EP 1 310 571 A2

	gcgccaccagg gctgcctgcc tccgttcccc ggggacgtct tcatgattcc tcagtacggg	1980
	tacctgactc tgaacaatgg cagtcaggcc gtgggcccgtt cctccttcta ctgcctggag	2040
5	tactttcctt ctcaaagtct gagaacgggc aacaactttg agttcagcta ccagtttgag	2100
	gacgtgcctt ttcacagcag ctacgcgcac agccaaagcc tggaccggct gatgaacccc	2160
	ctcatcgacc agtacctgta ctacctgtct cggactcagt ccacgggagg taccgcagga	2220
10	actcagcagt tgctattttc tcaggccggg cctaataaca tgcgggtca ggccaaaaac	2280
	tggctacccg ggcctgtcta ccggcagcaa cgcgtctcca cgacactgtc gcaaaataac	2340
	aacagcaact gtaaattccc gtgtcgctat ggcaaccac aaggacgacg aagagcgatt	2400
15	ttgcctggac cgggtgccacc aagtatcatc tgaatggcag agactctctg ttttcgctcc	2460
	agcggagtct taatgtttgg gaaacaggga gctggaaaag acaacgtgga ctatagcagc	2520
	gttatgctaa ccagtggagga agaaattaaa accaccaacc cagtggccac ggaacagtac	2580
20	ggcgtggtgg ccgataacct gcaacagcaa aacgccgctc ctattgtagg ggccgtcaac	2640
	agtcaaggag ccttacctgg catggtctgg cagaaccggg acgtgtacct gcagggtcct	2700
	atctgggcca agattcctca cacggacgga aactttcatc cctcgccgct gatgggaggc	2760
25	tttgactga aacacccgcc tctcagatc ctgattaaga atacacctgt tcccgcggat	2820
	cctccaacta cttcagtc agctaagctg gcgtcgttca tcacgcagta cagcaccgga	2880
	caggtcagcg tggaaattga atgggagctg cagaaagaaa acagcaaacg ctggaaccca	2940
30	gagattcaat acacttccaa ctactacaaa tctacaaatg tggacttcgc tgttaacaca	3000
	gatggcactt attctgagcc tcgccccatt ggcaccggtt acctcaccg taatctgtaa	3060
	ttgctcgta atcaataaac cggttgattc gtttcagttg aactttggtc tctgcgaagg	3120
35	gcgaattc	3128
	<210> 47	
	<211> 3128	
40	<212> DNA	
	<213> new AAV serotype, clone 44.5	
	<400> 47	
	gaattcgccc tttctacggc tgcgtaact ggaccaatga gaactttccc ttcaacgatt	60
45	gcgtcgacaa gatggtgata tgggtgggagg agggcaagat gacggccaag gtcgtggagt	120
	ccgccaaggc cattctcggc ggcagcaaag tgcgcggtga ccaaaagtgc aagtcgtccg	180
	cccagatcga cccaccccc gtgatcgtca cctccaacac caacatgtgc gccgtgattg	240
50	acgggaacag caccaccttc gagcaccagc agccgttgca ggaccggatg ttcaagtttg	300
	aactcaccg ccgtctggag cagactttg gcaaggtgac aaagcaggaa gtcagagagt	360
	tcttcgctg ggcgaggat cagtgaccg aggtggcgca cgagttctac gtcagaaagg	420
55	gtggagccaa caagagacc gcccccgatg acgcggataa aagcgagccc aagcgggcct	480

EP 1 310 571 A2

	gccccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccgggtg gactttgccc	540
	acagggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca	600
5	aaacatgcga gagaatgaat cagaatttca acatttgctt cagcacggg accagagact	660
	gttcagaatg tttccccggc gtgtcagaat ctcaaccggg tgtcagaaaa aagacgtatc	720
	ggaaactctg tgcgattcat catctgctgg ggccggcacc cgagattgct tgctcggcct	780
10	gcgatctggg caacgtggac ctagatgact gtgtttctga gcaataaatg acttaaacca	840
	ggtatggctg ccgatgggta tcttccagat tggctcgagg acaacctctc tgagggcatt	900
	cgcgagtggg gggacttgaa acctggagcc ccgaaacca aagccaacca gcaaaagcag	960
15	gacgacggcc ggggtctggg gcttctggc tacaagtacc tcggaccctt caacggactc	1020
	gacaaggggg agcccgtaaa cgcgcgggac gcagcggccc tcgagcacga caaggcctac	1080
	gaccagcagc tcaaagcggg tgacaatccg tacctgcggg ataaccacgc cgacgccgag	1140
20	tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcgggag agcagtcttc	1200
	caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct	1260
	cctggaaaga agagaccggg agagccatca cccagcgtt ctccagactc ctctacgggc	1320
25	atcggcaaga aaggccagca gcccgcgaaa aagagactca actttgggca gactggcgac	1380
	tcagagtcag tgcccgaccc tcaaccaatc ggagaacccc ccgcaggccc ctctggtctg	1440
	ggatctggta caatggctgc aggcggtggc gctccaatgg cagacaataa cgaaggcgcc	1500
30	gacggagtgg gtagttctc aggaatttg cattgcgatt ccacatggct gggcgacaga	1560
	gtcatcacca ccagcaccgg aacctgggccc ctccccacct acaacaacca cctctacaag	1620
	caaattctca acgggacttc gggaggaagc accaacgaca acacctactt cggctacagc	1680
35	acccctggg ggtattttga ctttaacaga ttccactgcc acttctcacc acgtgactgg	1740
	cagcgactca tcaacaacaa ctggggattc cggcccaaga gacccaactt caagctcttc	1800
	aacatccagg tcaaggaggc cagcgagaat gaaggcacca agaccatcgc caataacctt	1860
40	accagcacga ttcagggtctt tacggactcg gaataccagc tcccgtacgt cctcggctct	1920
	gcgcaccagg gctgcctgcc tccgttcccg gcggacgtct tcatgattcc tcagtacggg	1980
	tacctgactc tgaacaatgg cagtcaggcc gtgggcccgtt cctccttcta ctgcctggag	2040
45	tactttcctt ctcaaagtct gagaacgggc aacaactttg agttcagcta ccagtttgag	2100
	gacgtgcctt ttcacagcag ctacgcgcac agccaaagcc tggaccggct gatgaacccc	2160
	ctcatcgacc agtacctgta ctacctgtct cggactcagt ccacgggagg taccgcagga	2220
50	actcagcagt tgctattttc tcaggccggg cctaataaca tgtcggctca ggccaaaaac	2280
	tggctacccg ggccctgcta ccggcagcaa cgcgtctcca cgacactgtc gcaaaataac	2340
55	aacagcaact ttgcctggac cgggtgccacc aagtatcatc tgaatggcag agactctctg	2400

# EP 1 310 571 A2

	gtaaatcccg gtgtcgctat ggcaaccac aaggacgacg aagagcgatt tttccgctcc	2460
	agcggagtct taatgtttgg gaaacaggga gctggaaaag acaacgtgga ctatagcagc	2520
5	gttatgctaa ccagtgagga agaaattaaa accaccaacc cagtggccac agaacagtac	2580
	ggcgtggtgg ccgataacct gcaacagcaa aacgccgctc ctattgtagg ggccgtcaac	2640
	agtcaaggag ccttacctgg catggtctgg cagaaccggg acgtgtacct gcagggctct	2700
10	atctgggcca agattcctca caccggacgga aactttcatc cctcgccgct gatgggaggc	2760
	tttgactga aacaccgcc tcctcagatc ctgattaaga atacacctgt tcccgcggat	2820
	cctccaacta ccttcagtca agctaagctg gcgtcggttca tcacgcagta cagcaccgga	2880
15	caggtcagcg tggaaattga atgggagctg cagaaagaaa acagcaaacg ctggaaccca	2940
	gagattcaat acacttccaa ctactacaaa tctacaaatg tggactttgc tgttaacaca	3000
	gatggcactt attctgagcc tcgccccatt ggcaccggtt acctcaccg taatctgtaa	3060
20	ttgcttggtta atcaataaac cggttgattc gtttcagttg aactttggtc tctgcgaagg	3120
	gcgaattc	3128
25	<210> 48 <211> 1933 <212> DNA <213> new AAV serotype, clone 223.10	
30	<220> <221> misc_feature <222> (1302)..(1302) <223> can be a, c, g or t	
	<400> 48	
35	caaggcctac gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc	60
	cgacgccgag tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg	120
	agcagtcttc caggccaaaa agcgggttct cgaacctctt ggtctggttg agacgccagc	180
40	taagacggca cctggaaaga agcgaccggt agactcgcca gactccacct cgggcacg	240
	caagaaaggc cagcagcccg cgaaaaagag actcaacttt gggcagactg gcgactcaga	300
	gtcagtcccc gacctcaac caatcgga accaccagca ggccccctctg gtctgggac	360
45	tggtacaatg gctgcaggcg gtggcgacc aatggctgac aataacgagg gcgccgacgg	420
	agtgggtaat gcctcaggaa attggcattg cgattccaca tggtggggcg acagagtcac	480
	caccaccagc acccgaacct ggccccctgcc cacctacaac aaccacctct acaagcaa	540
50	ctccagtcag tcagcaggga gcaccaacga taacgtctat ttcggctaca gcacccctg	600
	ggggatattt gacttcaaca gattccattg ccacttctca ccacgtgact ggcagcgact	660
	tatcaacaac aactggggat tccggcccaa gaagctcaac ttcaagctct tcaacatcca	720
55	ggtcaaggag gtcacgacga atgacggtgt cacaaccatc gctaataacc ttaccagcac	780

# EP 1 310 571 A2

ggttcaggtc ttttcggact cggaatatca actgccgtac gtccctcggt ccgcgcacca 840  
 gggctgcctg cctccgttcc cggcagacgt gttcatgatt ccgcagtagc gatacctgac 900  
 5 tctgaacaat ggcagccaat cggtaggccg ttcctccttc tactgcctgg agtactttcc 960  
 ttctcagatg ctgagaacgg gcaacaactt cacccttagc tacaccttcg aggacgtgcc 1020  
 tttccacagc agctacggcg acagccagag tctggaccgg ctgatgaatc cctcatcga 1080  
 10 ccagtacctg tactacttgg ccagaacaca gagcaacgca ggaggtactg ctggcaatcg 1140  
 ggaactgcag ttttatcagg gcggacctac caccatggcc gaacaagcaa agaactggct 1200  
 gcccgacact tgcttcggcg aacagagagt atccaagacg ctggatcaaa ataacaacag 1260  
 15 caactttgcc tggactggtg ccacaaaata ccatttaaatt gnaagaaatt cattgggttaa 1320  
 tcccgggtgc gccatggcaa ccacaaagga cgacgaggaa cgcttcttcc cttcgagcgg 1380  
 agttctaatt tttggcaaaa ctggagcagc taataaaact acattagaaa acgtgctcat 1440  
 20 gacaaatgaa gaagaaattc gtctaccaa cccggtagct accgaggaat acgggattgt 1500  
 aagcagcaac ttgcaggcgg ctagcaccgc agccagaca caagttgtta acaaccaggg 1560  
 agccttacct ggcatggtct ggcagaaccg ggacgtgtac ctgcaaggtc ccatttgggc 1620  
 25 caagattcct cacacggacg gcaactttca cccgtctcct ctaatgggtg gctttggact 1680  
 gaaacacccg cctccccaga tcctgatcaa aaacacaccg gtacctgcta atcctccaga 1740  
 agtggtttact cctgccaagt ttgtctcctt catcacgcag tacagcaccg ggcaagtcag 1800  
 30 cgttgagatc gagtgggagc tgcagaaaga gaacagcaag cgctggaacc cagagattca 1860  
 gtacacctcc aactttgaca aacagactgg agtggacttt gctgttgaca gccagggtgt 1920  
 ttactctgag cct 1933  
 35  
 <210> 49  
 <211> 1933  
 <212> DNA  
 <213> new AAV serotype, clone 223.2  
 40  
 <400> 49  
 caaggcctac gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc 60  
 cgacgccgag tttcaggagt gtcttcaaga agatacgtct tttgggggca acctcgggcg 120  
 45 agcagtcttc caggccaaaa agcgggttct cgaacctctt ggtctggttg agacgccagc 180  
 taagacggca cctggaaaga agcgaccggt agactcgcca gactccacct cgggcacatcg 240  
 caagaaaggc cagcagcccg cgaaaaagag actcaacttt gggcagactg gcgactcaga 300  
 50 gtcagtcccc gaccctcaac caatcggaga accaccagca ggccctctct gtctgggac 360  
 tgggtacaatg gttgcaggcg gtggcgacc aatggctgac aataacgagg gcgccgacgg 420  
 agtgggtaat gcctcaggaa attggcattg cgattccaca tggctgggcg acagagtcac 480  
 55 caccaccagc acccgaacct ggccctgcc cacctacaac aaccacctct acaagcaa 540

# EP 1 310 571 A2

	ctccagtcag tcagcagggg gcaccaacga taacgtctat ttcggctaca gcacccctg	600
	ggggtatttt gacttcaaca gattccattg ccactttctca ccacgtgact ggcagcgact	660
5	tatcaacaac aactggggat tccggcccaa gaagctcaac ttcaagctct tcaacatcca	720
	ggtcaaggag gtcacgacga atgacgggtg cacaaccatc gctaataacc ttaccagcac	780
	ggttcaggtc ttttcggact cggaaatca actgcccgtac gtcctcggct ccgcgcacca	840
10	gggctgcctg cctccgttcc cggcagacgt gttcatgatt ccgcagtagc gatacctgac	900
	tctgaacaat ggcagccaat cggtaggcg ttcctccttc tactgcctgg agtactttcc	960
	ttctcagatg ctgagaacgg gcaacaactt cacctttagc tacaccttcg aggacgtgcc	1020
15	tttccacagc agctacgcgc acagccagag tctggaccgg ctgatgaatc ccctcatcga	1080
	ccagtacctg tactacttgg ccagaacaca gagcaacgca ggaggtactg ctggcaatcg	1140
	ggaactgcag ttttatcagg gcggacctac caccatggcc gaacaagcaa agaactggct	1200
20	gcccgacct tgcttcgggc aacagagagt atccaagacg ctggatcaaa ataacaacag	1260
	caactttgcc tggactggtg ccacaaaata ccatttaaatt ggaagaaatt cattggttaa	1320
	tcccgggtgc gccatggcaa cccacaagga cgacgaggaa cgcttctccc cttcgagcgg	1380
25	agttctaatt tttggcaaaa ctggagcagc taataaaact acattagaaa acgtgctcat	1440
	gacaaatgaa gaagaaattc gtcctaccaa cccggtagct accgaggaat acgggattgt	1500
	aagcagcaac ttgcaggcgg ctagcacccg agcccagaca caagtgttta acaaccaggg	1560
30	agccttacct ggcatggtct ggacgaaccg ggacgtgtac ctgcaaggtc ccatttgggc	1620
	caagattcct cacacggacg gcaactttca cccgtctcct ctaatgggtg gctttggact	1680
	gaaacacccg cctccccaga tcctgatcaa aaacacgccg gtacctgcta atcctccaga	1740
35	agtgtttact cctgccaaagt ttgcttcctt catcacgcag tacagcaccg ggcaagtacg	1800
	cgttgagatc gagtgggagc tgcagaaaga gaacagcaag cgctggaacc cagagattca	1860
	gtacacctcc aactttgaca aacagactgg agtggacttt gctgttgaca gccagggtgt	1920
40	ttactctgag cct	1933
	<210> 50	
	<211> 1933	
45	<212> DNA	
	<213> new AAV serotype, clone 223.4	
	<400> 50	
	caaggcctac gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc	60
50	cgacgccgag tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg	120
	agcagtcttc caggccaaaa agcgggttct cgaacctctt ggtctggttg agacgccagc	180
	taagacggca cctggaaaga agcgaccggt agactcgcca gactccacct cgggcacgcg	240
55	caagaaaggc cagcagcccc cgaaaaagag actcaacttt gggcagactg gcgactcaga	300

# EP 1 310 571 A2

	gccagtgccc gaccctcaac caatcggaga accaccagca ggcccctctg gtctgggac	360
	tggtacaatg gctgcaggcg gtggcgacc aatggctgac aataacgagg gcgccgacgg	420
5	agtgggtaat gcctcaggaa attggcattg cgattccaca cggctgggcg acagagtcac	480
	caccaccagc acccgaacct gggccctgcc cacctacaac aaccacctct acaagcaa	540
	ctccagtcag tcagcaggga gcaccaacga taacgtctat ttgggtaca gcacccctg	600
10	gggggtatatt gacttcaaca gattccattg ccacttctca ccacgtgact ggcagcgact	660
	tatcaacaac aactggggat tccggcccaa gaagctcaac ttcaagctct tcaacatcca	720
	ggtcaaggag gtcacgacga atgacggcg cacaaccatc gctaataacc ttaccagcac	780
15	ggttcagggtc ttttcggact cggaatatca actgccgtac gtccctcggct ccgcgacca	840
	gggctgcctg cctccgttcc cggcagacgt gttcatgatt ccgcagtagc gatacctgac	900
	tctgaacaat ggcagccaat cggtagggcg ttcctccttc tactgcctgg agtactttcc	960
20	ttctcagatg ctgagaacgg gcaacaactt cacctttagc tacaccttcg aggacgtgcc	1020
	tttccacagc agctacggc acagccagag tctggggccgg ctgatgaatc cctcatcga	1080
	ccagtacctg tactacttgg ccagaacaca gagcaacgca ggaggtagtg ctggcaatcg	1140
25	ggaactgcag ttttatcagg gcggacctac caccatggcc gaacaagcaa agaactggct	1200
	gcccggacct tgcttccggc aacagagagt atccaagacg ctggatcaaa ataacaacag	1260
	caactttgcc tggactgggt ccacaaaata ccatttaaat ggaagaaatt cattgggtta	1320
30	tcccgggtgtc gccatggcaa cccacaagga cgacgaggaa cgcttcttcc cttcgagcgg	1380
	agttctaatt tttggcaaaa ctggagcagc taataaaaact acattagaaa acgtgctcat	1440
	gacaaatgaa gaagaaatc gtcctacca cccggtagct accgaggaat acgggattgt	1500
35	aagcagcaac ttgcaggcgg ctagcaccgc agcccagaca caagttgtta acaaccaggg	1560
	agccttacct ggcatggtct ggcagaaccg ggacgtgtac ctgcaaggct ccatttgggc	1620
	caagattcct cacacggacg gcaactttca cccgtctcct ctaatgggtg gctttggact	1680
40	gaaacacccg cctccccaga tctgatcaa aaacacaccg gtacctgcta atcctccaga	1740
	agtgtttact cctgccaagt ttgcttcctt catcacgcag tacagcaccg ggcaagtcag	1800
	cgttgagatc gaatgggagc tgcagaaaga gaacagcaag cgctggaacc cagagattca	1860
45	gtacacctcc aactttgaca aacagactgg agtggacttt gctgttgaca gccagggtgt	1920
	ttactctgag cct	1933
50	<210> 51	
	<211> 1933	
	<212> DNA	
	<213> new AAV serotype, clone 223.5	
	<400> 51	
55	caaggcctac gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc	60

EP 1 310 571 A2

	cgacgccgag	tttcaggagc	gtcttcaaga	agatacgtct	tttgggggca	acctcgggcg	120
	agcagtcttc	caggccaaaa	agcgggttct	cgaacctctt	ggtctggttg	agacgccagc	180
5	taagacggca	cctggaaaga	agcgaccggt	agactcgcca	gactccacct	cgggcacatcg	240
	caagaaaggc	cagcagcccg	cgaaaaagag	actcaacttt	gggcagactg	gcgactcaga	300
	gccagtcccc	gacctcaac	caatcgga	accaccagca	ggccccctctg	gtctgggatc	360
10	tggtacaatg	gctgcaggcg	gtggcgacc	aatggctgac	aataacgagg	gcgccgacgg	420
	agtgggtaat	gcctcaggaa	attggcattg	cgattccaca	cggtcgggcg	acagagtcac	480
	caccaccagc	acccgaacct	gggccctgcc	cacctacaac	aaccacctct	acaagcaa	540
15	ctccagtcag	tcagcaggga	gcaccaacga	taacgtctat	ttcggctaca	gcacccccctg	600
	ggggatattt	gacttcaaca	gattccattg	ccacttctca	ccacgtgact	ggcagcgact	660
	tatcaacaac	aactggggat	tccggcccaa	gaagctcaac	ttcaagctct	tcaacatcca	720
20	ggtcaaggag	gtcacgacga	atgacggcgt	cacaaccatc	gctaataacc	ttaccagcac	780
	ggttcaggtc	ttttcggact	cggaatatca	actgccgtac	gtcctcggct	ccgcgcacca	840
	gggctgcctg	cctccgttcc	cggcagacgt	gttcattgatt	ccgcagtagc	gatacctgac	900
25	tctgaacaat	ggcagccaat	cggtaggcgg	ttcctccttc	tactgcctgg	agtactttcc	960
	ttctcagatg	ctgagaacgg	gcaacaactt	cacctttagc	tacaccttcg	aggacgtgcc	1020
	tttccacagc	agctacgcgc	acagccagag	tctggggccg	ctgatgaatc	ccctcatcga	1080
30	ccagtacctg	tactacttgg	ccagaacaca	gagcaacgca	ggaggtagct	ctggcaatcg	1140
	ggaactgcag	ttttatcagg	gcggacctac	caccatggcc	gaacaagcaa	agaactggct	1200
	gccccgacct	tgcttccggc	aacagagagt	atccaagacg	ctggatcaaa	ataacaacag	1260
35	caactttgcc	tggactggtg	ccacaaaata	ccatttaa	ggaagaaatt	cattgggttaa	1320
	tccccgtgtc	gccatggcaa	cccacaagga	cgacgaggaa	cgttcttcc	cttcgagcgg	1380
	agttctaatt	tttggcaaaa	ctggagcagc	taataaaact	acattagaaa	acgtgctcat	1440
40	gacaaatgaa	gaagaaattc	gtcctaccaa	cccggtagct	accgaggaat	acgggattgt	1500
	aagcagcaac	ttgcaggcgg	ctagcaccgc	agcccagaca	caagttgtta	acaaccaggg	1560
	agccttacct	ggcatggtct	ggcagaaccg	ggacgtgtac	ctgcaaggtc	ccatttgggc	1620
45	caagattcct	cacacggacg	gcaactttca	cccgtctcct	ctaattgggtg	gctttggact	1680
	gaaacacccg	cctccccaga	tcctgatcaa	aaacacaccg	gtacctgcta	atcctccaga	1740
	agtgtttact	cctgccaa	ttgcttcctt	catcacgcag	tacagcaccg	ggcaagtcag	1800
50	cgttgagatc	gaatgggagc	tgcaaaaaga	gaacagcaag	cgttggaacc	cagagattca	1860
	gtacacctcc	aactttgaca	aacagactgg	agtggacttt	gctgttgaca	gccagggtgt	1920
55	ttactctgag	cct					1933

# EP 1 310 571 A2

<210> 52  
 <211> 1933  
 <212> DNA  
 <213> new AAV serotype, clone 223.6

5

<400> 52  
 caaggcctac gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc 60  
 cgacgccgag tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg 120  
 10 agcagtcttc caggccaaaa agcgggttct cgaacctctt ggtctggttg agacgccagc 180  
 taagacggca cctggaaga agcgaccggt agactcgcca gactccacct cgggcatcgg 240  
 caagaaaggc cagcagcccg cgaaaaagag actcaacttt gggcagactg gcgactcaga 300  
 15 gtcagtcccc gacctcaac caatcgga accaccagca ggccctctg gtctgggatc 360  
 tgggtacaatg gctgcaggcg gtggcgacc aatggctgac aatagcgagg gcgccgacgg 420  
 agtgggtaat gcctcaggaa attggcattg cgattccaca tggctgggcg acagagtcac 480  
 20 caccaccagc acccgaacct gggccctgcc cacctacaac aaccacctct acaagcaaata 540  
 ctccagtcag tcagcaggga gcaccaacga taacgtctat ttcggctaca gcacccctg 600  
 ggggtatattt gacttcaaca gattccattg ccacttctca ccacgtgact ggcagcgact 660  
 25 tatcaacaac aactggggat tccggcccaa gaagctcaac ttcaagctct tcaacatcca 720  
 ggtcaaggag gtcacgacga atgacggtgt cacaaccatc gctaataacc ttaccagcac 780  
 ggttcaggtc ttttcggact cggaatatca actgccgtac gtctcggct ccgcgcacca 840  
 30 gggctgcctg cctccgttcc cggcagacgt gttcatgatt ccgcagtagc gatacctgac 900  
 tctgaacaat ggcagccaat cggtaggccc ttcctccttc tactgcctgg agtactttcc 960  
 ttctcagatg ctgagaacgg gcaacaactt cacctttagc tacaccttcg aggacgtgcc 1020  
 35 tttccacagc agctacgagc acagccagag tctggaccgg ctgatgaatc ccctcatcga 1080  
 ccagtacctg tactacttgg ccagaacaca gagcaacgca ggaggtagtg ctggcaatcg 1140  
 ggaactgcag ttttatcagg gcggacctac caccatggcc gaacaagcaa agaactggct 1200  
 40 gcccgacct tgcttccggc aacagagagt atccaagacg ctggatcaaa ataacaacag 1260  
 caactttgcc tggactggtg ccacaaaata ccatttaaata ggaagaaatt cattgggtta 1320  
 tcccggtgtc gccatggcaa ccacaaagga cgacgaggaa cgcttcttcc ctccgagcgg 1380  
 45 agtttctaatt tttggcaaaa ctggagcagc taataaaact acattagaaa acgtgctcat 1440  
 gacaaatgaa gaagaaattc gtcctacca cccggtagct accgaggaat acgggattgt 1500  
 aagcagcaac ttgcaggcgg ctagcaccgc agccagaca caagttgtta acaaccaggg 1560  
 50 agccttacct ggcatggtct ggcagaaccg ggacgtgtac ctgcaaggtc ccatttgggc 1620  
 caagattcct cacacggagc gcaactttca ccgctctcct ctaatgggtg gctttggact 1680  
 gaaacacccg cctccccaga tctgatcaa aaacacaccg gtacctgcta atcctccaga 1740  
 55

# EP 1 310 571 A2

	agtgtttact cctgccaaagc ttgcttcctt catcacgcag tacagcacccg ggcaagtcag	1800
5	cgttgagatc gagtgggagc tgcagaaaga gaacagcaag cgctggaacc cagagattca	1860
	gtacacctcc aactttgaca aacagactgg agtggacttt gctgttgaca gccaggggtgt	1920
	ttactctgag cct	1933
10	<210> 53 <211> 1933 <212> DNA <213> new AAV serotype, clone 223.7	
15	<400> 53 caaggcctac gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc	60
	cgacgccgag ttctcaggagc gtcttcaaga agatacgtct tttgggggca acctcggggcg	120
	agcagtcttc caggccaaaa agcgggttct cgaacctctt ggtctggttg agacgccagc	180
20	taagacggca cctggaaaga agcgaccggt agactcgcca gactccacct cgggcatcgg	240
	caagaaaggc cagcagcccg cgaaaaagag actcaacttt gggcagactg gcgactcaga	300
	gtcagtcccc gacctcaac caatcggaga accaccagca ggccccctctg gtctgggac	360
25	tggtacaatg gctgcaggcg gtggcgccacc aatggctgac aataacgagg gcgccgacgg	420
	agtgggtaat gcctcaggaa attggcattg cgattccaca tggctggggc acagagtcac	480
	caccaccagc acccgaaacct gggccctgcc cacctacaac aaccacctct acaagcaaat	540
30	ctccagtcag tcagcaggga gcaccaacga taacgtctat ttcggctaca gccccctg	600
	ggggtatattt gacttcaaca gattccattg ccacttctca ccacgtgact ggcagcgact	660
	tatcaacaac aactggggat tccggcccaa gaagctcaac ttcaagctct tcaacatcca	720
35	ggtcaaggag gtcacgacga atgacggcgt cacaaccatc gctaataacc ttaccagcac	780
	ggttcaggtc ttttcggacc cggaatatca actgccgtac gtctctcggct ccgcgcacca	840
	gggctgcctg cctccgttcc cggcagacgt gttcatgatt ccgcagtacg gatacctgac	900
40	tctgaacaat ggcagccaat cggtaggcg ttctctcttc tactgcctgg agtactttcc	960
	ttctcagatg ctgagaacgg gcaacaactt cacctttagc tacaccttcg aggacgtgcc	1020
	tttccacagc agctacgcgc acagccagag tctggaccgg ctgatgaatc ccctcatcga	1080
45	ccagttacctg tactacttgg ccagaacaca gagcaacgca ggaggtactg ctggcaatcg	1140
	ggaactgcag ttttatcagg gcggacctac caccatggcc gaacaagcaa agaactggct	1200
	gcccggacct tgcttcgggc aacagagagt atccaagacg ctggatcaaa ataacaacag	1260
50	caactttgcc tggactggtg ccacaaaata ccatttaaatt ggaagaaatt cattgggttaa	1320
	tcccgggtgc gccatggcaa cccacaagga cgacgaggaa cgcttcttcc cttcgagcgg	1380
	agttctaatt tttggcaaaa ctggagcagc taataaaaact acattagaaa acgtgctcat	1440
55	gacaaatgaa gaagaaattc gtcctaccaa cccggtagct accgaggaat acgggattgt	1500

# EP 1 310 571 A2

	aagcagcaac ttgcaggcgg ctagcacccg agcccagaca caagttgtta acaaccaggg	1560
	agccttacct ggcatgggtc ggcagaaccg ggacgtgtac ctgcaaggtc ccatttgggc	1620
5	caagattcct cacacggacg gcaactttca cccgtctcct ctaatgggtg gctttggact	1680
	gaaacacccg cctccccaga tctgatcaa aaacacaccg gtacctgcta atcctccaga	1740
	agtgtttact cctgccaaga ttgttctctt catcacgcag tacagcaccc ggcaagtcag	1800
10	cgttgagatc gagtgggagc tgcagaaaga gaacagcaag cgctggaacc cagagattca	1860
	gtacacctcc aactttgaca aacagactgg agtggacttt gctgttgaca gccagggtgt	1920
	ttactctgag cct	1933
15		
	<210> 54	
	<211> 3123	
	<212> DNA	
	<213> new AAV serotype, clone A3.4	
20		
	<400> 54	
	gaattcgccc tttctacggc tgcgtcaact ggaccaatga aaactttccc ttcaacgatt	60
	gcgctcgacaa gatgggtgatc tgggtggagg agggaaagat gaccgccaag gtcgtggaat	120
25	ctgccaaagc cattctgggt ggaagcaagg ttcgtgtgga ccagaaatgc aagtcttcgg	180
	cccagatcga cccgactccg gtgattgtca cctctaacac caacatgtgc gccgtgattg	240
	acggaaactc gaccaccttc gagcaccagc agccgttgca agaccggatg ttcaaatttg	300
30	aacttacccg ccgtttggat catgactttg ggaaggtcac caagcaggaa gtcaaagact	360
	ttttccgggtg ggctcaagat cacgtgactg aggtggagca tgagttctac gtcaaaaagg	420
	gtggagccaa gaaaaggccc gccccgatg atgtatatat aaatgagccc aagcgggcgc	480
35	gcgagtcagt tgcgcagcca tcgacgtcag acgcggaagc ttcgataaac tacgcgggca	540
	ggtaccaaaa caaatgttct cgtcacgtgg gcatgaatct gatgctgttt ccctgtcgac	600
	aatgcgaaag aatgaatcag aattcaata tctgcttcac acacgggcaa aaagactgtt	660
40	tggaatgctt tcccgtgtca gaatctcaac ccgtttctgt cgtcagaaaa acgtatcaga	720
	aactttgtta cattcatcat atcatgggaa aagaaccaga cgctgcact gcctgcgacc	780
	tggtaaatgt ggacttggat gactgtatct ctgagcaata aatgacttaa atcaggtatg	840
45	gctgctgacg gttatcttcc agattggctc gaggacactc tctctgaagg aatcagacag	900
	tggtggaagc tcaaacctgg cccaccaccg ccgaaacctt accaacaaca ccgggacgac	960
	agtaggggtc ttgtgcttcc tgggtacaag tacctcggac ccttcaacgg actcgacaaa	1020
50	ggagagcccg tcaacgaggc agacgccgcg gccctcgagc acgacaaagc ctacgaccac	1080
	cagctcaagc aaggggacaa cccgtacctc aaatacaacc acgcggacgc tgaatttcag	1140
	gagcgtcttc aagaagatac gtctttcggg ggcaacctcg ggcgagcagt cttccaggcc	1200
55	aaaaagaggg tactcgagcc tcttggctctg gttgaggaag ctgttaagac ggctcctgga	1260

EP 1 310 571 A2

	aaaaagagac ctatagagca gtctcctgca gaaccggact cttcctcggg catcggcgaa	1320
	tcaggccagc agcccgttaa gaaaagactc aatttttggtc agactggcga cacagagtca	1380
5	gtcccagacc ctcaaccaat cggagaaccc cccgcagccc cctctggtgt gggatctaat	1440
	acaatggctt caggcgggtg ggcaccaatg gcagacgata acgaaggcgc cgacggagt	1500
	ggtaattcct cgggaaattg gcattgcgat tccacatgga tgggcgacag agttatcacc	1560
10	accagcacia gaacctgggc cctccccacc tacaataatc acctctacia gcaaatctcc	1620
	agcgaatcgg gagccaccaa cgacaaccac tacttcggct acagcaccac ctgggggtat	1680
	tttgacttta acagattcca ctgtcacttc tcaccacgtg actggcagcg actcatcaac	1740
15	aacaactggg gatttagacc caagaaactc aatttcaagc tcttcaacat ccaagtcaag	1800
	gaggtcacgc agaattgatg aaccacgacc atcgccaata accttaccag cacgggtcag	1860
	gtcttcacag actctgagta ccagctgccc tacgtcctcg gttcgggtca ccagggctgc	1920
20	cttcgccgt tcccagcaga cgtcttcacg attcctcagt acggctactt gactctgaac	1980
	aatggcagcc aagcggtagg acgtttctca ttctactgtc tagagtattt tccctctcag	2040
	atgctgagga cgggaaacaa cttcaccttc agctacactt ttgaagacgt gcctttccac	2100
25	agcagctacg cgcacagcca gagtctggat cggctgatga atcctctcat tgaccagtac	2160
	ctgtattacc tgagcaaaac tcaggttaca agtggaacaa cgcagcaatc gagactgcag	2220
	ttcagccaag ctgggcctag ctccatggct cagcaggcca aaaactggct accgggaccc	2280
30	agctaccgac agcagcgaat gtctaagacg gctaattgaca acaacaacag tgaatttgct	2340
	tggactgcag ccaccaaata ttacctgaat ggaagaaatt ctctggtcaa tcccgggccc	2400
	ccaatggcca gtcacaagga cgatgaggaa agtattttcc ccatgcacgg aaatctcatc	2460
35	tttgaaaaac aaggcacagg aactaccaat gtggacattg aatcagtgtt tattacagac	2520
	gaagaagaaa tcagaacaac taatcctgtg gctacagAAC aatacggaca ggttgccacc	2580
	aaccatcaga gtcaggacac cacagcttcc tatggaagtg tggacagcca gggaatctta	2640
40	cctggaatgg tgtggcagga ccgcgatgtc tatcttcaag gtcccatttg ggccaaaact	2700
	cctcacacgg acggacactt tcctccttct ccgctcatgg gaggttttg actgaaacac	2760
	cctcctcccc agatcctgat caaaaacaca cctgtgccag cgaatccgc gaccactttc	2820
45	actcctggaa agtttgcttc gttcattacc cagtattcca ccggacaggt cagcgtggaa	2880
	atagagtggg agctgcagaa agaaaacagc aaacgctgga accagaaat tcagtacacc	2940
	tccaactaca acaagtccgt gaatgtggag ttaccgtgg acgcaaacgg tgtttattct	3000
50	gaaccccgcc ctattggcac tcgttacctt acccggaact tgtaatttcc tgttaatgaa	3060
	taaaccgatt tatgcgtttc agttgaactt tggctctctgc gaagggcgaa ttcgcggccg	3120
55	cta	3123

# EP 1 310 571 A2

<210> 55  
 <211> 3113  
 <212> DNA  
 <213> new AAV serotype, clone A3.5

<400> 55  
 gaattcgccc tttctacggc tgcgtcaact ggaccaatga aaactttccc ttcaacgatt 60  
 gcgtcgacaa gatggtgatc tgggtgggagg agggaaagat gaccgccaag gtcgtggaat 120  
 ctgccaaaagc cattctgggt ggaagcaagg ttcgtgtgga ccagaaatgc aagtcttcgg 180  
 cccagatcga cccgactccg gtgattgtca cctctaacac caacatgtgc gccgtgattg 240  
 acggaaactc gaccaccttc gagcaccagc agccgttgca agaccggatg ttcaaatttg 300  
 aacttaccgg ccgtttggat catgactttg ggaaggtcac caagcaggaa gtcaaagact 360  
 ttttcgggtg ggctcaagat cacgtgactg aggtggagca tgagttctac gtcaaaaagg 420  
 gtggagccaa gaaaaggccc gccccgatg atgtatatat aaatgagccc aagcggggcg 480  
 gcgagtcagt tgcgcagcca tcgacgtcag acgcggaagc ttcgataaac tacgcgga 540  
 ggtacaaaaa caaatgttct cgtcacgtgg gcatgaatct gatgctgttt cctgtcgac 600  
 aatgcgaaag aatgaatcag aattcaaata tctgcttcac acacgggcaa aaagactgtt 660  
 tggaatgctt tcccggtgca gaatctcaac ccgttcctgt cgtcagaaaa acgtatcaga 720  
 aactttgtta cattcatcat atcatgggaa aagtaccaga cgcctgcact gcctgcgacc 780  
 tggtaaagtgt ggacttggat gactgtatct ctgagcaata aatgacttaa atcaggtatg 840  
 gctgctgacg gttatcttcc agattggctc gaggacactc tctctgaagg aatcagacag 900  
 tgggtgaagc tcaaacctgg cccaccaccg ccgaaacctt accaacaaca ccgggacgac 960  
 agtaggggtc ttgtgcttcc tgggtacaag tacctcggac ctttcaacgg actcgacaaa 1020  
 ggagagccgg tcaacgaggc agacgccggc gccctcgagc acgacaaagc ctacgaccac 1080  
 cagctcaagc aaggggacaa cccgtacctc aaatacaacc acgcggacgc tgaatttcag 1140  
 gagcgtcttc aagaagatac gtctttcggg ggcaacctcg ggcgagcagt cttccaggcc 1200  
 aaaaagaggg tactcgagcc tcttgggtctg gttgaggaag ctgttaagac ggctcctgga 1260  
 aaaaagagac ctatagagca gtctcctgca gaaccggact cttcctcggg catcggcaaa 1320  
 tcaggccagc agcccgctaa gaaaagactc aattttggtc agactggcga cacagagtca 1380  
 gtcccagacc ctcaaccaat cggagaacct cccgcagccc cctctggtgt gggatcta 1440  
 acaatggctt caggcgggtg ggcaccaatg gcagacaata acgaaggcgc cgacggagt 1500  
 ggtaattcct cgggaaattg gcattgcgat tccacatgga tgggagacag agttatcacc 1560  
 accagcacia gaacctgggc cctccccacc tacaataatc acctctacaa gcaaattctc 1620  
 agcgaatcgg gagccaccaa cgacaaccac tacttcgggt acagcacccc ctgggggtat 1680  
 tttgacttta acagattcca ctgtcacttc tcaccacgtg actggcagcg actcatcaat 1740

# EP 1 310 571 A2

	aacaactggg gatttagacc caagaaactc aatttcaagc tcttcaacat ccaagtcaag	1800
	gaggtcacgc agaattgatgg aaccacgacc atcgccaata accttaccag cagggtgcag	1860
5	gtcttcacag actctgagta ccagctgccc tacgtcctcg gttcggctca ccagggctgc	1920
	cttccgccgt tcccagcaga cgtcttcatg attcctcagt acggctactt gactctgaac	1980
	aatggcagcc aagcggtagg acgtttctca ttctactgtc tagagtatth tccctctcag	2040
10	atgctgagga cgggaaacaa cttcaccttc agctacactt ttgaagacgt gcctttccac	2100
	agcagctacg cgcacagcca gagtctggat cggctgatga atcctctcat tgaccagtac	2160
	ctgtattacc tgagcaaaac tcaggggtaca agtggaacaa cgcagcaatc gagactgcag	2220
15	ttcaaccaag ctgggcctag ctccatggct cagcaggcca aaaactggct accgggaccc	2280
	agctaccgac agcagcgaat gtctaagacg gctaatagaca acaacaacag tgaatttgct	2340
	tggactgcag ccaccaata ttacccgaat ggaagaaatt ctctggtcaa tccggggccc	2400
20	ccaatggcca gtcacaagga cgatgaggaa aagtatttcc ccatgcacgg aaatctcatc	2460
	tttggaanaa aaggcacagg aactaccaat gtggacattg aatcagtgtc tattacagac	2520
	gaagaagaaa tcagaacgac taatcctgtg gctacagaac aatacggaca ggttgccacc	2580
25	aaccgtcaga gtcagaacac cacagcttcc tatggaagtg tggacagcca gggaatctta	2640
	cctggaatgg tgtggcagga ccgcgatgtc tatcttcaag gtcccatttg ggccaaaact	2700
	cctcacacgg acggacactt tcatccttct ccgctcatgg gaggccttgg actgaaacac	2760
30	cctcctcccc agatcctgat caaaaacaca cctgtgccag cgaatccgcg gaccacttcc	2820
	actcctggaa agtttgcttc gtccattacc cagtattcca ccggacaggt cagcgtggaa	2880
	atagagtggg agctgcagaa agaaaacagc aaacgctgga acccggaatc tcagtacacc	2940
35	tccaactaca acaagtcggg gaatgtggag ttaccgtgg acgcaaacgg tgtttattct	3000
	gaaccccgcc ctattggcac tcgttacctt acccggaact tgtaatttcc tgttaatgaa	3060
40	taaaccgatt tatgcgtttc agttgaactt tggctctctgc gaagggcgaa ttc	3113
	<210> 56	
	<211> 3122	
	<212> DNA	
	<213> new AAV serotype, clone A3.7	
45	<400> 56	
	agcggccgcg aattcgccct ttctacggct gcgtcaactg gaccaatgaa aactttccct	60
	tcaacgattg cgtcgacaag atggtgatct ggtgggagga gggaaagatg accgccaagg	120
50	tcgtggaatc tgccaaagcc attctgggtg gaagcaaggt tcgtgtggac cagaaatgca	180
	ggtcttcggc ccagatcgac ccgactccgg tgattgtcac ctctaacacc aacatgtgcg	240
	ccgtgattga cggaaactcg accaccttcg agcaccagca gccgttgcaa gaccggatgt	300
55	tcaaatttga acttaccgcg cgtttggatc atgactttgg gaaggtcacc aagcaggaag	360

EP 1 310 571 A2

	tcaaagactt tttccggtgg gctcaagatc acgtgactga ggtggagcat gagttctacg	420
	tcaaaaaggg tggagccaag aaaaggcccg ccccgatga tgtatatata aatgagccca	480
5	agcgggcgcg cgagtcagtt gcgcagccat cgacgtcaga cgcggaagct tcgataaact	540
	acgcggacag gtacccaaac aaatgtttctc gtcacgtggg catgaatctg atgctgtttc	600
10	cctgtcgaca atgcgaaaga atgaatcaga attcaaatat ctgcttcaca cacgggcaaa	660
	aagactgttt ggaatgcttt cccgtgtcag aatctcaacc cgtttctgtc gtcagaaaaa	720
	cgtatcagaa actttgttac attcatcata tcatgggaaa agtaccagac gcctgcactg	780
15	cctgcgacct ggtaaatgtg gacttggatg actgtatttc tgagcaataa atgacttaaa	840
	tcaggatatgg ctgctgacgg ttatcttcca gattggctcg aggacactct ctctgaagga	900
	atcagacagt ggtggaagct caaacctggc ccaccaccgc cgaaacctaa ccaacaacac	960
20	cgggacgaca gtaggggtct tgtgcttctt ggttacaagt acctcggacc cttcaacgga	1020
	ctcgacaaaag gagagccggc caacgaggca gacgccggcg ccctcgagca cgacaaagcc	1080
	tacgaccacc agctcaagca aggggacaac ccgtacctca aatacaacca cgcggacgct	1140
25	gaatttcagg agcgtcttca agaagatacg tctttcgggg gcaacctcgg gcgagcagtc	1200
	ttccaggcca aaaagagggt actcgagcct cttggtctgg ttgaggaagc tgttaagacg	1260
	gctcctggaa aaaagagacc tatagagcag tctcctgcag aaccggactc ttctcgggc	1320
30	atcggcaaatt caggccagca gcccgctaag aaaagactca attttggta gactggcgac	1380
	acagagtcag tcccagacct tcaaccaatc ggagaacccc ccgcagcccc ctctggtgtg	1440
	ggatctaata caatggcttc aggcggtggg gcaccaatgg cagacaataa cgaaggcgcc	1500
35	gacggagtgg gtaattcttc gggaaattgg cattgcgatt ccacatggat gggcgacaga	1560
	gttatcacca ccagcacaag aacctgggac ctccccacct acaataatcg cctctacaag	1620
	caaatctcca gcgaatcggg agccaccaac gacaacct acttcggcta cagcaccccc	1680
40	tgggggtatt ttgactttaa cagattccac tgtcacttct caccacgtga ctggcagcga	1740
	ctcatcaaca acaactgggg atttagacct aagaaactca atttcaagct cttcaacatc	1800
	caagtcaagg aggtcacgca gaatgatgga accacgacca tcgccaataa ccttaccagc	1860
45	acggtgcagg tcttcacaga ctctgagtac cagctgccct acgtcctcgg ttcggctcac	1920
	cagggctgcc ttccgcggtt cccagcagac gtcttcatga ttctctagta cggctacttg	1980
	actctgaaca atggcagcca agcggtagga cgttcttcat tctactgtct agagtatttt	2040
50	ccctctcaga tgctgaggac gggaaacaac ttcaccttca gctacacttt tgaagacgtg	2100
	cctttccaca gcagctacgc gcacagccag agtctggatc ggctgatgaa tctctcatt	2160
	gaccagtacc tgtattacct gagcaaaact caggggtacaa gtggaacaac gcagcaatcg	2220
55	agactgcagt tcagccaagc tgggcctagc tccatggctc agcaggccaa aaactggcta	2280

# EP 1 310 571 A2

	ccgggaccca gctaccgaca gcagcgaatg tctaagacgg ctaatgacaa caacaacagt	2340
	gaatttgctt ggactgcagc caccaaatat tacctgaatg gaagaaattc tctgggtcaat	2400
5	cccgggcccc caatggccag tcacaaggac gatgaggaaa agtatttccc catgcacgga	2460
	aatctcatct ttggaaaaca aggcacagga actaccaatg tggacattga atcagtgtctt	2520
10	attacagacg aagaagaaat cagaacaact aatcctgtgg ctacagaaca atacggacag	2580
	gttgccacca accatcagag tcagaacacc acagcttcct atggaagtgt ggacagccag	2640
	ggaatcttac ctggaatggg gtggcaggac cgcgatgtct atcttcaagg tcccatttgg	2700
15	gccaaaactc ctcacacgga cggacacttt catccttctc cgctcatggg aggctttgga	2760
	ctgaaacacc ctctcccca gatcctgatc aaaaacacac ctgtgccagc gaatcccgcg	2820
	accactttca ctcttggaat gtttgcctcg ttcattacc agtattccac cggacaggtc	2880
20	agcgtggaat tagagtggga gctgcagaaa gaaaacagca aacgctggaa ccagaaaatt	2940
	cagtacacct ccaactacaa caagtcggtg aatgtggagt ttaccgtgga cgcaaacggt	3000
	gtttattctg aaccccgccc tattggcact cgttacetta cccggaactt gtaatttctt	3060
25	gttaatgaat aaaccgattt atgcgtttca gttgaacttt ggtctctgcg aagggcgaat	3120
	tc	3122
30	<210> 57 <211> 3123 <212> DNA <213> new AAV serotype, clone A3.3	
	<400> 57	
	gaattcgccc tttctacggc tgcgtcaact ggaccaatga aaactttccc ttcaacgatt	60
35	gcgtcgacaa gatggtgatc tgggtgggagg agggaaagat gaccgccaag gtcgtggaat	120
	ctgccaaagc cattctgggt ggaggcaagg ttcgtgtgga ccagaaatgc aagtcttcgg	180
	cccagatcga cccgactccg gtgattgtca cctctaacac caacatgtgc gccgtgattg	240
40	acggaaactc gaccaccttc gagcaccagc agccgttgca agaccggatg ttcaaatttg	300
	aacttaccgg ccgtttggat catgactttg ggaagggtcac caagcaggaa gtcaaagact	360
	ttttccggtg ggctcaagat cacgtgactg aggtggagca tgagttctac gtcaaaaagg	420
45	gtggagccaa gaaaaggccc gccccgatg atgtatatat aaatgagccc aagcgggcgc	480
	gcgagtcagt tgcgcagcca tcgacgtcag acgcggaagc ttcgataaac tacgcggaca	540
	ggtaccaaaa caaatgttct cgtcacgtgg gcatgaatct gatgctgttt ccctgtcgac	600
50	aatgcgaaag aatgaatcag aattcaaata tctgcttcac acacgggcaa aaagactggt	660
	tggaatgctt tcccgtgtca gaatctcaac ccgtttctgt cgtcagaaaa acgtatcaga	720
	aactttgtta cattcatcat atcatgggaa aagtaccaga cgctgcact gcctgcgacc	780
55	tggtaaatgt ggacttggat gactgtattt ctgagcaata aatgacttaa atcaggtatg	840

EP 1 310 571 A2

	gctgctgacg gttatcttcc agattggctc gaggacactc tctctgaagg aatcagacag	900
	tgggtggaagc tcaaacctgg cccaccaccg ccgaaacctc accaacaaca ccgggacgac	960
5	agtaggggtc ttgtgcttcc tgggtacaag tacctcggac ccttcaacgg actcgacaaa	1020
	ggagagccgg tcaacgaggc agacgccgcg gccctcgagc acgacaaagc ctacgaccac	1080
10	cagctcaagc aaggggacaa cccgtacctc aaatacaacc acgcgagcgc tgaatttcag	1140
	gagcgtcttc aagaagatac gtctttcggg ggcaacctcg ggcgagcagt cttccaggcc	1200
	aaaaagaggg tactcgagcc tcttggctctg gttgaggaag ctgttaagac ggctcctgga	1260
15	aaaaagagac ctatagagca gtctcctgca gaaccggact cttcctcggg catcggcaaa	1320
	tcaggccagc agcccgctaa gaaaagactc aatttttggtc agactggcga cacagagtca	1380
	gtcccaggcc ctcaaccaat cggagaacct cccgcagccc cctctggtgt gggatctaat	1440
20	acaatggctt caggcggtgg ggcaccaatg gcagacaata acgaaggcgc cgacggagtg	1500
	ggtaattcct cgggaaattg gcattgcgat tccacatgga tgggcgacag agttatcacc	1560
	accagcacia gaacctgggc cctccccacc tacaataatc acctctacaa gcaaatctcc	1620
25	agcgaatcgg gagccaccaa cgacaaccac tacttcggct acagcaccac ctgggggtat	1680
	tttgacttta acagattoca ctgtcacttc tcaccacgtg actggcagcg actcatcaac	1740
	aacaactggg gatttagacc caagaaactc aatttcaagc tcttcaacat ccaagtcaag	1800
30	gaggtcacgc agaatgatgg aaccacgacc atcgccaata accttaccag cgcggtgcag	1860
	gtcttcacag actctgagta ccagctgccc tacgtcctcg gttcggctca ccagggctgc	1920
	cttcgcgcgt tcccagcaga cgtcttcatg attcctcagt acggctactt gactctgaac	1980
35	aatggcagcc aagcggtagg acgttcttca ttctactgtc tagagtattt tccctctcag	2040
	atgctgagga cgggaaacaa cttcaccttc agctacactt ttgaagacgt gcctttccac	2100
	agcagctacg cgacacagca gagtctggat cggctgatga atcctctcat tgaccagtac	2160
40	ctgtattacc tgagcaaaac tcagggtaga agtgaacaa cgcagcaatc gagactgcag	2220
	ttcagccaag ctgggcctag ctccatggct cagcaggcca aaaactggct accgggaccc	2280
	agctaccgac agcagcgaat gtctaagacg gctaatagaca acaacaacag tgaatttgct	2340
45	tggactgcag ccaccaata ttacctgaat ggaagaaatt ctctggtcaa tcccgggccc	2400
	ccagtggcca gtcacaagga cgatgaggaa aagtatttcc ccatgcacgg aaatctcatc	2460
	tttgaaaaac aaggcacagg aactaccaat gtggacattg aatcagtgtc tattacagac	2520
50	gaagaagaaa tcagaacaac taatcctgtg gctacagaac aatacggaca ggttgccacc	2580
	aaccatcaga gtcagaacac cacagcttcc tatggaagtg tggacagcca gggaatctta	2640
	cctggaatgg tgtggcagga ccgcgatgtc tatcttcaag gtcccatttg ggccaaaact	2700
55	cctcacacgg acggacactt tcctccttct ccgctcatgg gaggctttgg actgaaacac	2760

# EP 1 310 571 A2

	cctcctcccc agatcctgat caaaaacaca cctgtgccag cgaatcccg caccactttc	2820
	actcctggaa agtttgcttc gttcattacc cagtattcca cctgacaggt cagcgtggaa	2880
5	atagagtggg agctgcagaa agaaaacagc aaacgctgga acccagaaat tcagtacacc	2940
	tccaactaca acaagtccgt gaatgtggag tttaccgtgg acgcaaacgg tgttttattct	3000
	gaaccccgcc ctattggcac tcgttacctt acccggaact tgtaatttcc tgттаатgaa	3060
10	taagccgatt tatgcgtttc agttgaactt tggctctctgc gaagggcgaa ttcgтттааа	3120
	cct	3123
15	<210> 58 <211> 2969 <212> DNA <213> new AAV serotype, clone 42.12	
	<400> 58	
20	gaattcgccc tttctacggc tgcgtaact ggaccaatga gaactttccc ttcaacgatt	60
	gcgtcgacaa gatggtgatc tgggtgggagg agggcaagat gacggccaag gtcgtggagt	120
	ccgccaaggc cattctcggc ggcagcaagg tgcgcggtga ccaaagtgc aagtcgtccg	180
25	cccagatcga cccaccccc gtgatcgtca cctccaacac caacatgtgc gccgtgattg	240
	acgggaacag caccaccttc gagcaccagc agccgttaca agaccgatg ttcaaatttg	300
	aactcaccgc ccgtctggag cagcactttg gcaaggtgac aaagcaggaa gtcaaagagt	360
30	tcttccgctg ggcgagcat cactgtaccg aggtggcgca tgagttctac gtcagaaagg	420
	gtggagccaa caagagaccc gccccgatg acgcggataa aagcgagccc aagcgggcct	480
	gcccctcagt cgggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg	540
35	acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg ttccctgca	600
	agacatgca gagaatgaat cagaatttca acatttgctt cagcacggg accagagact	660
	gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc	720
40	ggaaactctg tgccattcat catctgctgg ggcggtctc cgagattgct tgctcggcct	780
	gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaacca	840
	ggtatggctg ccgatggtta tttccagat tggctcgagg acaacctctc tgagggcatc	900
45	cgcgagtggg gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag	960
	gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc	1020
	gacaagggag agccggtcaa cgaggcagac gccgcggccc tcgagcacga caaggcctac	1080
50	gacaagcagc tcgagcaggg ggacaacccg tacctcaagt acaaccacgc cgacgccgag	1140
	tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg agcagtcctc	1200
	caggccaaga agcgggttct cgaacctctc ggtctgggtg aggaaggcgc taagacggct	1260
55	cctggaaaga agagaccggt agagccatca cccagcgtt ctccagactc ctctacgggc	1320

# EP 1 310 571 A2

5 atcggcaaga caggccagca gcccgcaaaa aagagactca actttgggca gactggcgac 1380  
 tcagagtcag tgcccgaccc tcaaccaatc ggagaacccc ccgcaggccc ctctgggtctg 1440  
 ggatctggta caatggctgc aggcgggtggc gctccaatgg cagacaataa cgaaggcgcc 1500  
 10 gacggagtgg gtagttcctc aggaaattgg cattgcgatt ccacatggct gggcgacaga 1560  
 gtcacaccca ccagcaccgc aacctggggc ctccccacct acaacaacca cctctacaag 1620  
 caaatctcca acgggacatc gggaggaagc accaacgaca acacctactt cggctacagc 1680  
 accccctggg ggtattttga ctttaacaga ttccactgcc acttctcacc acgtgactgg 1740  
 15 cagcgactca tcaacaacaa ctggggatcc cggcccaaga gactcaactt caagctcttc 1800  
 aacatccagg tcaaggaggt cagcgagaat gaaggcacca agaccatcgc caataacctt 1860  
 accagcacga ttcaggtctt tacggactcg gaataccagc tcccgtagct cctcggtctt 1920  
 20 gcgcaccagg gctgcctgcc tccgttcccg gcggacgtct tcattgattcc tcagtacggg 1980  
 tacctgactc tgaacaacgg cagtcaggcc gtgggcctgt cctccttcta ctgcctggag 2040  
 tactttcctt ctcaaagtct gagaacgggc aacaactttg agttcagcta ccagtttgag 2100  
 25 gacgtgcctt ttcacagcag ctacgcgcac agccaaagcc tggaccggct gacgaacccc 2160  
 ctcatcgacc agtacctgta ctacctggcc cggaccaga gcactacggg gtccacaagg 2220  
 gggctgcagt tccatcaggc tgggccaac accatggccg agcaatcaaa gaactggctg 2280  
 30 cccggacctt gttatcgga cagagactg tcaaaaaaca tagacagcaa caacaacagt 2340  
 aactttgccc ggaccggggc cactaaatac catctgaatg gtagaaatc attaaccaac 2400  
 ccgggcgtag ccatggccac caacaaggac gacgaggacc agttctttcc catcaacgga 2460  
 35 gtgctggttt ttggcaaac gggggctgcc aacaagacaa cgctggaaaa cgtgctaattg 2520  
 accagcgagg aggagatcaa aaccaccaat ccgctggcta cagaagaata cgggtgtggtc 2580  
 tccagcaacc tgcaatcgtc tacggccgga cccagacac agactgtcaa cagccagggg 2640  
 40 gctctgcccg gcatggtctg gcagaaccgg gacgtgtacc tgcagggtcc catctgggcc 2700  
 aaaattcctc acacggacgg caactttcac ccgtctcccc tgatggggcg atttgactc 2760  
 aaacacccgc ctctcaaat tctcalcaag talacttcca actactacaa atctacaaat 2820  
 45 gtggactttg ctgtcaatac tgagggtact tattcagagc ctgcgcccat tggcaccctg 2880  
 tacctcaccg gtaacctgta attgcctgtt aatcaataaa ccggttaatt cgtttcagtt 2940  
 gaacttttgt ctctgcgaag ggcgaattc 2969  
 50  
 <210> 59  
 <211> 3129  
 <212> DNA  
 <213> new AAV serotype, clone 44.2  
 55  
 <400> 59  
 gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60

# EP 1 310 571 A2

5           gcgtcgacaa gatggtgata tgggtgggagg aggggcaagat gacggccaag gtcgtggagt   120  
           ccgccaaaggc catttctcggc ggcagcaaaag tgcgcgtgga ccaaaagtgc aagtcgtccg   180  
           cccagatcga cccaccccc gtgatcgtea cctccaacac caacatgtgc gccgtgattg   240  
 10          acgggaacag caccaccttc gagcaccagc agccgttgca ggaccggatg ttcaagtttg   300  
           aactcaccgg ccgtctggag cactgactttg gcaagggtgac aaagcaggaa gtcagagagt   360  
           tcttccgctg gggcgaggat cactgaccg aggtggcgca cgagttctac gtcagaaagg   420  
           gtggagccaa caagagaccc gccccgatg acgcgataa aagcgagccc aagcgggccc   480  
 15          gccccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg   540  
           acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca   600  
           aaacatgcga gagaatgaat cagaatttca acatttgctt cactgacggg accagagact   660  
 20          gttcagaatg tttccccggc gtgtcagaat ctcaaccggc cgtcagaaaa aagacgtatc   720  
           ggaaactctg tgcgattcat catctgctgg gggcgggcac ccgagattgc ttgctcggcc   780  
           tgcgatctgg tcaacgtgga ctagatgac tgtgtttctg agcaataaat gacttaaacc   840  
 25          aggtatggct gccgatggtt atcttccaga ttggctcgag gacaacctct ctgagggcat   900  
           tcgcgagtgg tgggacttga aacctggagc ccgaaaaccc aaagccaacc agcaaaagca   960  
           ggacgacggc cggggtctgg tgcttcctgg ctacaagtac ctcgaccctt tcaacggact   1020  
 30          cgacaagggg gagcccgta acgcgcgga cgcagcgcc ctcgagcacg acaaggccta   1080  
           cgaccagcag ctcaaagcgg gtgacaatcc gtacctgcgg tataaccacg ccgacgccga   1140  
           gtttcaggag cgtctgcaag aagatacgtc ttttgggggc aacctcgggc gagcagtctt   1200  
 35          ccaggccaag aagcgggttc tcgaacctct cggctctggtt gaggaaggcg ctaagacggc   1260  
           tcctggaaaag aagagaccgg tagagccatc accccagcgt tctccagact cctctacggg   1320  
           catcggaag aaagccagc agcccgcaa aaagagactc aactttgggc agactggcga   1380  
 40          ctcagagtca gtgcccgacc ctcaaccaat cggagaaccc ccgcaggcc cctctggtct   1440  
           gggatctggt acaatggctg caggcgggtg cgctccaatg gcagacaata acgaaggcgc   1500  
           cgacggagtg ggtagttcct caggaaattg gcattgcgat tccacatggc tgggcgacag   1560  
 45          agtcatacc accagcaccg gaacctgggc cctccccacc tacaacaacc acctctacaa   1620  
           gcaaatctcc aacgggactt cgggaggaag caccaacgac aacacctact tcggctacag   1680  
           caccctctgg gggatatttg actttaacag attccactgc cacttctcac cactgactg   1740  
 50          gcagcgactc atcaacaaca actggggatt cggcccaaag agactcaact tcaagctctt   1800  
           caacatccag gtcaaggagg tcacgcagaa tgaaggcacc aagaccatcg ccaataacct   1860  
           taccagcacg attcaggtct ttacggactc ggaataccag ctcccgtagc tcctcggctc   1920  
 55          tgcgaccag ggctgcctgc ctccgttccc ggcggacgtc ttcattgattc ctcagtacgg   1980

**EP 1 310 571 A2**

5	gtacctgact	ctgaacaatg	gcagtcaggc	cgtagggccgt	tccctcttct	actgcctgga	2040
	gtacttttct	tctcaaatgc	tgagaacggg	caacaacttt	gagttcagct	accagtttga	2100
	ggacgtgcct	tttcacagca	gctacgcgca	cagccaaagc	ctggaccggc	tgatgaaccc	2160
10	cctcatcgac	cagtacctgt	actacctgtc	tcggactcag	tccacgggag	gtaccgcagg	2220
	aactcagcag	ttgctatttt	ctcaggccgg	gcctaataac	atgtcggctc	aggccaaaaa	2280
	ctggctaccc	gggccttget	accggcagca	acgcgtctcc	acgacactgt	cgaaaataa	2340
15	caacagcaac	tttgccctgga	ccggtgccac	caagtatcat	ctgaatggca	gagactctct	2400
	ggtaaattccc	ggtgtcgcta	tggcaaccca	caaggacgac	gaagagcgat	tttttcgctc	2460
	cagcggagtc	ttaatgtttg	ggaaaacagg	agctggaaaa	gacaacgtgg	actatagcag	2520
20	cgttatgcta	accagtgagg	aagaaattaa	aaccaccaac	ccagtggcca	cagaacagta	2580
	cggcgtggtg	gccgataacc	tgcaacagca	aaacgccgct	cctattgtag	gggccgtcaa	2640
	cagtcaagga	gccttacctg	gcattggtctg	gcagaaccgg	gacgtgtacc	tgcagggtcc	2700
25	tatctggggc	aagattcctc	acacggacgg	aaactttcat	ccctcgccgc	tgatgggagg	2760
	ctttggactg	aaacacccgc	ctcctcagat	cctgattaag	aatacacctg	ttcccgcgga	2820
	tcctccaact	accttcagtc	aagctaagct	ggcgtcgttc	atcacgcagt	acagcaccgg	2880
30	acaggtcagc	gtggaaattg	aatgggagct	gcagaaagaa	aacagcaaac	gctggaaccc	2940
	agagattcaa	tacacttcca	actactacaa	atctacaaat	gtggactttg	ctgttaacac	3000
	agatggcact	tattctgagc	ctcgccccat	cggcaccgct	tacctaccc	gtaatctgta	3060
35	attgcttggt	aatcaataaa	ccggttgatt	cgtttcagtt	gaactttggg	ctctgcgaag	3120
	ggcgaattc						3129

```
<210> 60
<211> 733
<212> PRT
<213> capsid protein of AAV serotype, clone C1VP1
<400> 60
```

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

# EP 1 310 571 A2

5 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
85 90 95

10 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
100 105 110

15 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
130 135 140

20 Pro Leu Glu Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly Lys  
145 150 155 160

25 Lys Gly Lys Gln Pro Ala Lys Lys Arg Leu Asn Phe Glu Glu Asp Thr  
165 170 175

Gly Ala Gly Asp Gly Pro Pro Glu Gly Ser Asp Thr Ser Ala Met Ser  
180 185 190

30 Ser Asp Ile Glu Met Arg Ala Ala Pro Gly Gly Asn Ala Val Asp Ala  
195 200 205

35 Gly Gln Gly Ser Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys  
210 215 220

Asp Ser Thr Trp Ser Glu Gly Lys Val Thr Thr Thr Ser Thr Arg Thr  
225 230 235 240

40 Trp Val Leu Pro Thr Tyr Asn Asn His Leu Tyr Leu Arg Leu Gly Thr  
245 250 255

45 Thr Ser Asn Ser Asn Thr Tyr Asn Gly Phe Ser Thr Pro Trp Gly Tyr  
260 265 270

Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln  
275 280 285

50 Arg Leu Ile Asn Asn Asn Trp Gly Leu Arg Pro Lys Ala Met Arg Val  
290 295 300

55 Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Thr Ser Asn Gly Glu  
305 310 315 320

# EP 1 310 571 A2

5	Thr Thr Val Ala Asn Asn Leu Thr Ser Thr Val Gln Ile Phe Ala Asp	325	330	335
	Ser Ser Tyr Glu Leu Pro Tyr Val Met Asp Ala Gly Gln Glu Gly Ser	340	345	350
10	Leu Ser Pro Phe Pro Asn Asp Val Phe Met Val Pro Gln Tyr Gly Tyr	355	360	365
	Cys Gly Ile Val Thr Gly Glu Asn Gln Asn Gln Thr Asp Arg Asn Ala	370	375	380
15	Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly Asn	385	390	395
	Asn Phe Glu Met Ala Tyr Asn Phe Gly Lys Val Pro Phe His Ser Met	405	410	415
20	Tyr Ala Tyr Ser Gln Ser Pro Asp Arg Leu Met Asn Pro Leu Leu Asp	420	425	430
25	Gln Tyr Leu Trp His Leu Gln Ser Thr Thr Ser Gly Glu Thr Leu Asn	435	440	445
	Gln Gly Asn Ala Ala Thr Thr Phe Gly Lys Ile Arg Ser Gly Asp Phe	450	455	460
30	Ala Phe Tyr Arg Lys Asn Trp Leu Pro Gly Pro Cys Val Lys Gln Gln	465	470	475
35	Arg Leu Ser Lys Thr Ala Ser Gln Asn Tyr Lys Ile Pro Ala Ser Gly	485	490	495
	Gly Asn Ala Leu Leu Lys Tyr Asp Thr His Tyr Thr Leu Asn Asn Arg	500	505	510
40	Trp Ser Asn Ile Ala Pro Gly Pro Pro Met Ala Thr Ala Gly Pro Ser	515	520	525
45	Asp Gly Asp Phe Ser Asn Ala Gln Leu Ile Phe Pro Gly Pro Ser Val	530	535	540
	Thr Gly Asn Thr Thr Thr Ser Ala Asn Asn Leu Leu Phe Thr Ser Glu	545	550	555
50	Glu Glu Ile Ala Ala Thr Asn Pro Arg Asp Thr Asp Met Phe Gly Gln	565	570	575
55				

# EP 1 310 571 A2

5 Ile Ala Asp Asn Asn Gln Asn Ala Thr Thr Ala Pro Ile Thr Gly Asn  
 580 585 590  
 Val Thr Ala Met Gly Val Leu Pro Gly Met Val Trp Gln Asn Arg Asp  
 595 600 605  
 10 Ile Tyr Tyr Gln Gly Pro Ile Trp Ala Lys Ile Pro His Ala Asp Gly  
 610 615 620  
 15 His Phe His Pro Ser Pro Leu Ile Gly Gly Phe Gly Leu Lys His Pro  
 625 630 635 640  
 Pro Pro Gln Ile Phe Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ala  
 645 650 655  
 20 Thr Thr Phe Thr Ala Ala Arg Val Asp Ser Phe Ile Thr Gln Tyr Ser  
 660 665 670  
 25 Thr Gly Gln Val Ala Val Gln Ile Glu Trp Glu Ile Glu Lys Glu Arg  
 675 680 685  
 Ser Lys Arg Trp Asn Pro Glu Val Gln Phe Thr Ser Asn Tyr Gly Asn  
 690 695 700  
 30 Gln Ser Ser Met Leu Trp Ala Pro Asp Thr Thr Gly Lys Tyr Thr Glu  
 705 710 715 720  
 35 Pro Arg Val Ile Gly Ser Arg Tyr Leu Thr Asn His Leu  
 725 730  
 <210> 61  
 <211> 733  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone C2VP1  
 40 <400> 61  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 45 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Leu  
 20 25 30  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 50 Gly Tyr Lys Tyr Leu Gly Pro Phe His Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 55 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80

# EP 1 310 571 A2

5 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 10 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 15 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 Pro Leu Glu Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly Lys  
 145 150 155 160  
 20 Lys Gly Lys Gln Pro Ala Lys Lys Arg Leu Asn Phe Glu Glu Asp Thr  
 165 170 175  
 25 Gly Ala Gly Asp Gly Pro Pro Glu Gly Ser Asp Thr Ser Ala Met Ser  
 180 185 190  
 Ser Asp Ile Glu Met Arg Ala Ala Pro Gly Gly Asn Ala Val Asp Ala  
 195 200 205  
 30 Gly Gln Gly Ser Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys  
 210 215 220  
 35 Asp Ser Thr Trp Ser Glu Gly Lys Val Thr Thr Thr Ser Thr Arg Thr  
 225 230 235 240  
 Trp Val Leu Pro Thr Tyr Asn Asn His Leu Tyr Leu Arg Leu Gly Thr  
 245 250 255  
 40 Thr Ser Asn Ser Asn Thr Tyr Asn Gly Phe Ser Thr Pro Trp Gly Tyr  
 260 265 270  
 45 Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln  
 275 280 285  
 Arg Leu Ile Asn Asn Asn Trp Gly Leu Arg Pro Lys Ala Met Arg Val  
 290 295 300  
 50 Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Thr Ser Asn Gly Glu  
 305 310 315 320  
 55 Thr Thr Val Ala Asn Asn Leu Thr Ser Thr Val Gln Ile Phe Ala Asp  
 325 330 335

# EP 1 310 571 A2

5	Ser Ser Tyr Glu Leu Pro Tyr Val Met Asp Ala Gly Gln Glu Gly Ser	340	345	350
	Leu Pro Pro Phe Pro Asn Asp Val Phe Met Val Pro Gln Tyr Gly Tyr	355	360	365
10	Cys Gly Ile Val Thr Gly Glu Asn Gln Asn Gln Thr Asp Arg Asn Ala	370	375	380
15	Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly Asn	385	390	395
	Asn Phe Glu Met Ala Tyr Asn Phe Glu Lys Val Pro Phe His Ser Met	405	410	415
20	Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro Leu Leu Asp	420	425	430
25	Gln Tyr Leu Trp His Leu Gln Ser Thr Thr Ser Gly Glu Thr Leu Asn	435	440	445
	Gln Gly Asn Ala Ala Thr Thr Phe Gly Lys Ile Arg Ser Gly Asp Phe	450	455	460
30	Ala Phe Tyr Arg Lys Asn Trp Leu Pro Gly Pro Cys Val Lys Gln Gln	465	470	475
35	Arg Phe Ser Lys Thr Ala Ser Gln Asn Tyr Lys Ile Pro Ala Ser Gly	485	490	495
	Gly Asn Ala Leu Leu Lys Tyr Asp Thr His Tyr Thr Leu Asn Asn Arg	500	505	510
40	Trp Ser Asn Ile Ala Pro Gly Pro Pro Met Ala Thr Ala Gly Pro Ser	515	520	525
45	Asp Gly Asp Phe Ser Asn Ala Gln Leu Ile Phe Pro Gly Pro Ser Val	530	535	540
	Thr Gly Asn Thr Thr Thr Ser Ala Asn Asn Leu Leu Phe Thr Ser Glu	545	550	555
50	Gly Glu Ile Ala Ala Thr Asn Pro Arg Asp Thr Asp Met Phe Gly Gln	565	570	575
55	Ile Ala Asp Asn Asn Gln Asn Ala Thr Thr Ala Pro Ile Thr Gly Asn	580	585	590

# EP 1 310 571 A2

5 Val Thr Ala Met Gly Val Leu Pro Gly Met Val Trp Gln Asn Arg Asp  
 595 600 605  
 Ile Tyr Tyr Gln Gly Pro Ile Trp Ala Lys Ile Pro His Ala Asp Gly  
 610 615 620  
 10 His Phe His Pro Ser Pro Leu Ile Gly Gly Phe Gly Leu Lys His Pro  
 625 630 635 640  
 15 Pro Pro Gln Ile Phe Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ala  
 645 650 655  
 Thr Thr Phe Thr Ala Ala Arg Val Asp Ser Phe Ile Thr Gln Tyr Ser  
 660 665 670  
 20 Thr Gly Gln Val Ala Val Gln Ile Glu Trp Glu Ile Glu Lys Glu Arg  
 675 680 685  
 25 Ser Lys Arg Arg Asn Pro Glu Val Gln Phe Thr Ser Asn Tyr Gly Asn  
 690 695 700  
 Gln Ser Ser Met Leu Trp Ala Pro Asp Thr Thr Gly Lys Tyr Thr Glu  
 705 710 715 720  
 30 Pro Arg Val Ile Gly Ser Arg Tyr Leu Thr Asn His Leu  
 725 730  
 <210> 62  
 <211> 733  
 35 <212> PRT  
 <213> capsid protein of AAV serotype, clone C5VP1@2  
 <400> 62  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 40 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 45 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 Gly Tyr Glu Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 50 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 55 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95

# EP 1 310 571 A2

5 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 10 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 15 Pro Leu Glu Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly Lys  
 145 150 155 160  
 Lys Gly Lys Gln Pro Ala Lys Lys Arg Leu Asn Phe Glu Glu Asp Thr  
 165 170 175  
 20 Gly Ala Gly Asp Gly Pro Pro Glu Gly Ser Asp Thr Ser Ala Met Ser  
 180 185 190  
 25 Ser Asp Ile Glu Met Arg Ala Ala Pro Gly Gly Asn Ala Val Asp Ala  
 195 200 205  
 Gly Gln Gly Ser Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys  
 210 215 220  
 30 Asp Ser Thr Trp Ser Glu Gly Lys Val Thr Thr Thr Ser Thr Arg Thr  
 225 230 235 240  
 35 Trp Val Leu Pro Thr Tyr Asn Asn His Leu Tyr Leu Arg Leu Gly Thr  
 245 250 255  
 Thr Ser Asn Ser Asn Thr Tyr Asn Gly Phe Ser Thr Pro Trp Gly Tyr  
 260 265 270  
 40 Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln  
 275 280 285  
 45 Arg Leu Ile Asn Asn Asn Trp Gly Leu Arg Pro Lys Ala Met Arg Val  
 290 295 300  
 Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Thr Ser Asn Gly Glu  
 305 310 315 320  
 50 Thr Thr Val Ala Asn Asn Leu Thr Ser Thr Val Gln Ile Phe Ala Asp  
 325 330 335  
 55 Ser Ser Tyr Glu Leu Pro Tyr Val Met Asp Ala Gly Gln Glu Gly Ser  
 340 345 350

EP 1 310 571 A2

5 Leu Pro Pro Phe Pro Asn Asp Val Phe Met Val Pro Gln Tyr Gly Tyr  
355 360 365

Cys Gly Ile Val Thr Gly Glu Asn Gln Asn Gln Thr Asp Arg Asn Ala  
370 375 380

10 Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly Asn  
385 390 395 400

15 Asn Phe Glu Thr Ala Tyr Asn Phe Glu Lys Val Pro Phe His Ser Met  
405 410 415

Tyr Ala His Ser Gln Ser Leu Asp Gly Leu Met Asn Pro Leu Leu Asp  
420 425 430

20 Gln Tyr Leu Trp His Leu Gln Ser Thr Thr Ser Gly Glu Thr Leu Asn  
435 440 445

25 Gln Gly Asn Ala Ala Thr Thr Phe Gly Lys Ile Arg Ser Gly Asp Phe  
450 455 460

Ala Phe Tyr Arg Lys Asn Trp Leu Pro Gly Pro Cys Val Lys Gln Gln  
465 470 475 480

30 Arg Phe Ser Lys Thr Ala Ser Gln Asn Tyr Lys Ile Pro Ala Ser Gly  
485 490 495

35 Gly Asn Ala Leu Leu Lys Tyr Asp Thr His Tyr Thr Leu Asn Asn Arg  
500 505 510

Trp Ser Asn Ile Ala Pro Gly Pro Pro Met Ala Thr Ala Gly Pro Ser  
515 520 525

40 Asp Gly Asp Phe Ser Asn Ala Gln Leu Ile Phe Pro Gly Pro Ser Val  
530 535 540

45 Thr Gly Asn Thr Thr Thr Ser Ala Asn Asn Leu Leu Phe Thr Ser Glu  
545 550 555 560

Glu Glu Ile Ala Ala Thr Asn Pro Arg Asp Thr Asp Met Phe Gly Gln  
565 570 575

50 Ile Ala Asp Asn Asn Gln Asn Ala Thr Thr Ala Pro Ile Thr Gly Asn  
580 585 590

55 Val Thr Ala Met Gly Val Leu Pro Gly Met Val Trp Gln Asn Arg Asp  
595 600 605

# EP 1 310 571 A2

5 Ile Tyr Tyr Gln Gly Pro Ile Trp Ala Lys Ile Pro His Ala Asp Gly  
 610 615 620  
 His Phe His Pro Ser Pro Leu Ile Gly Gly Phe Gly Leu Lys His Pro  
 625 630 635 640  
 10 Pro Pro Gln Ile Phe Ile Lys Asn Thr Pro Val Pro Ala Tyr Pro Ala  
 645 650 655  
 15 Thr Thr Phe Thr Ala Ala Arg Val Asp Ser Phe Ile Thr Gln Tyr Ser  
 660 665 670  
 Thr Gly Gln Val Ala Val Gln Ile Glu Trp Glu Ile Glu Lys Glu Arg  
 675 680 685  
 20 Ser Lys Arg Trp Asn Pro Glu Val Gln Phe Thr Ser Asn Cys Gly Asn  
 690 695 700  
 25 Gln Ser Ser Met Leu Trp Ala Pro Asp Thr Thr Gly Lys Tyr Thr Glu  
 705 710 715 720  
 Pro Arg Val Ile Gly Ser Arg Tyr Leu Thr Asn His Leu  
 725 730  
 30 <210> 63  
 <211> 734  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone AAV4VP1  
 <400> 63  
 35 Met Thr Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser Glu  
 1 5 10 15  
 Gly Val Arg Glu Trp Trp Ala Leu Gln Pro Gly Ala Pro Lys Pro Lys  
 20 25 30  
 40 Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro Gly  
 35 40 45  
 45 Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro Val  
 50 55 60  
 Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp Gln  
 65 70 75 80  
 50 Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp  
 85 90 95  
 55 Ala Glu Phe Gln Gln Arg Leu Gln Gly Asp Thr Ser Phe Gly Gly Asn  
 100 105 110

# EP 1 310 571 A2

5           Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro Leu  
                   115                           120                           125  
  
           Gly Leu Val Glu Gln Ala Gly Glu Thr Ala Pro Gly Lys Lys Arg Pro  
                   130                           135                           140  
 10  
           Leu Ile Glu Ser Pro Gln Gln Pro Asp Ser Ser Thr Gly Ile Gly Lys  
           145                           150                           155                           160  
  
           Lys Gly Lys Gln Pro Ala Lys Lys Lys Leu Val Phe Glu Asp Glu Thr  
                   165                           170                           175  
 15  
  
           Gly Ala Gly Asp Gly Pro Pro Glu Gly Ser Thr Ser Gly Ala Met Ser  
                   180                           185                           190  
 20  
  
           Asp Asp Ser Glu Met Arg Ala Ala Ala Gly Gly Ala Ala Val Glu Gly  
                   195                           200                           205  
 25  
  
           Gly Gln Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys  
                   210                           215                           220  
  
           Asp Ser Thr Trp Ser Glu Gly His Val Thr Thr Thr Ser Thr Arg Thr  
           225                           230                           235                           240  
 30  
  
           Trp Val Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Arg Leu Gly Glu  
                   245                           250                           255  
  
           Ser Leu Gln Ser Asn Thr Tyr Asn Gly Phe Ser Thr Pro Trp Gly Tyr  
                   260                           265                           270  
 35  
  
           Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln  
                   275                           280                           285  
 40  
  
           Arg Leu Ile Asn Asn Asn Trp Gly Met Arg Pro Lys Ala Met Arg Val  
                   290                           295                           300  
  
           Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Thr Ser Asn Gly Glu  
           305                           310                           315                           320  
 45  
  
           Thr Thr Val Ala Asn Asn Leu Thr Ser Thr Val Gln Ile Phe Ala Asp  
                   325                           330                           335  
 50  
  
           Ser Ser Tyr Glu Leu Pro Tyr Val Met Asp Ala Gly Gln Glu Gly Ser  
                   340                           345                           350  
  
           Leu Pro Pro Phe Pro Asn Asp Val Phe Met Val Pro Gln Tyr Gly Tyr  
                   355                           360                           365  
 55

# EP 1 310 571 A2

5 Cys Gly Leu Val Thr Gly Asn Thr Ser Gln Gln Gln Thr Asp Arg Asn  
370 375 380

Ala Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly  
385 390 395 400

10 Asn Asn Phe Glu Ile Thr Tyr Ser Phe Glu Lys Val Pro Phe His Ser  
405 410 415

15 Met Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro Leu Ile  
420 425 430

Asp Gln Tyr Leu Trp Gly Leu Gln Ser Thr Thr Thr Gly Thr Thr Leu  
435 440 445

20 Asn Ala Gly Thr Ala Thr Thr Asn Phe Thr Lys Leu Arg Pro Thr Asn  
450 455 460

25 Phe Ser Asn Phe Lys Lys Asn Trp Leu Pro Gly Pro Ser Ile Lys Gln  
465 470 475 480

Gln Gly Phe Ser Lys Thr Ala Asn Gln Asn Tyr Lys Ile Pro Ala Thr  
485 490 495

30 Gly Ser Asp Ser Leu Ile Lys Tyr Glu Thr His Ser Thr Leu Asp Gly  
500 505 510

35 Arg Trp Ser Ala Leu Thr Pro Gly Pro Pro Met Ala Thr Ala Gly Pro  
515 520 525

Ala Asp Ser Lys Phe Ser Asn Ser Gln Leu Ile Phe Ala Gly Pro Lys  
530 535 540

40 Gln Asn Gly Asn Thr Ala Thr Val Pro Gly Thr Leu Ile Phe Thr Ser  
545 550 555 560

45 Glu Glu Glu Leu Ala Ala Thr Asn Ala Thr Asp Thr Asp Met Trp Gly  
565 570 575

Asn Leu Pro Gly Gly Asp Gln Ser Asn Ser Asn Leu Pro Thr Val Asp  
580 585 590

50 Arg Leu Thr Ala Leu Gly Ala Val Pro Gly Met Val Trp Gln Asn Arg  
595 600 605

55 Asp Ile Tyr Tyr Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp  
610 615 620

# EP 1 310 571 A2

5 Gly His Phe His Pro Ser Pro Leu Ile Gly Gly Phe Gly Leu Lys His  
625 630 635 640

Pro Pro Pro Gln Ile Phe Ile Lys Asn Thr Pro Val Pro Ala Asn Pro  
645 650 655

10 Ala Thr Thr Phe Ser Ser Thr Pro Val Asn Ser Phe Ile Thr Gln Tyr  
660 665 670

15 Ser Thr Gly Gln Val Ser Val Gln Ile Asp Trp Glu Ile Gln Lys Glu  
675 680 685

Arg Ser Lys Arg Trp Asn Pro Glu Val Gln Phe Thr Ser Asn Tyr Gly  
690 695 700

20 Gln Gln Asn Ser Leu Leu Trp Ala Pro Asp Ala Ala Gly Lys Tyr Thr  
705 710 715 720

25 Glu Pro Arg Ala Ile Gly Thr Arg Tyr Leu Thr His His Leu  
725 730

<210> 64  
<211> 736  
<212> PRT  
<213> capsid protein of AAV serotype, clone AAV1

30 <400> 64  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

35 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

40 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

45 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
85 90 95

50 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
100 105 110

55 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

# EP 1 310 571 A2

5           Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
               130                               135                               140  
  
           Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly  
           145                               150                               155                               160  
 10  
           Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
                                   165                               170                               175  
  
           Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro  
                                   180                               185                               190  
 15  
           Ala Thr Pro Ala Ala Val Gly Pro Thr Thr Met Ala Ser Gly Gly Gly  
                                   195                               200                               205  
 20  
           Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala  
                                   210                               215                               220  
  
           Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile  
           225                               230                               235                               240  
 25  
           Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
                                   245                               250                               255  
 30  
           Tyr Lys Gln Ile Ser Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His  
                                   260                               265                               270  
  
           Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe  
                                   275                               280                               285  
 35  
           His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn  
                                   290                               295                               300  
 40  
           Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln  
                                   305                               310                               315                               320  
  
           Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn  
                                   325                               330                               335  
 45  
           Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro  
                                   340                               345                               350  
  
           Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala  
                                   355                               360                               365  
 50  
           Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly  
                                   370                               375                               380  
 55

# EP 1 310 571 A2

5 Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro  
 385 390 395 400  
  
 Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe  
 405 410 415  
 10  
 Glu Glu Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp  
 420 425 430  
  
 Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg  
 435 440 445  
 15  
 Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn Lys Asp Leu Leu Phe Ser  
 450 455 460  
 20  
 Arg Gly Ser Pro Ala Gly Met Ser Val Gln Pro Lys Asn Trp Leu Pro  
 465 470 475 480  
 25  
 Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Lys Thr Asp Asn  
 485 490 495  
 Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala Ser Lys Tyr Asn Leu Asn  
 500 505 510  
 30  
 Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His Lys  
 515 520 525  
 35  
 Asp Asp Glu Asp Lys Phe Phe Pro Met Ser Gly Val Met Ile Phe Gly  
 530 535 540  
 Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala Leu Asp Asn Val Met Ile  
 545 550 555 560  
 40  
 Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Arg  
 565 570 575  
 Phe Gly Thr Val Ala Val Asn Phe Gln Ser Ser Ser Thr Asp Pro Ala  
 580 585 590  
 45  
 Thr Gly Asp Val His Ala Met Gly Ala Leu Pro Gly Met Val Trp Gln  
 595 600 605  
 50  
 Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His  
 610 615 620  
 55  
 Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu  
 625 630 635 640

# EP 1 310 571 A2

5 Lys Asn Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala  
 645 650 655  
 Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys Phe Ala Ser Phe Ile Thr  
 660 665 670  
 10 Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln  
 675 680 685  
 15 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Val Gln Tyr Thr Ser Asn  
 690 695 700  
 Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr Val Asp Asn Asn Gly Leu  
 705 710 715 720  
 20 Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Pro Leu  
 725 730 735  
 <210> 65  
 <211> 736  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone AAV6VP1  
 <400> 65  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 30 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 35 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 40 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 45 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 50 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 55 Phe Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140

# EP 1 310 571 A2

5 Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly  
 145 150 155 160  
  
 Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
 165 170 175  
 10  
 Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro  
 180 185 190  
  
 Ala Thr Pro Ala Ala Val Gly Pro Thr Thr Met Ala Ser Gly Gly Gly  
 195 200 205  
 15  
 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala  
 210 215 220  
 20  
 Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile  
 225 230 235 240  
  
 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
 245 250 255  
 25  
 Tyr Lys Gln Ile Ser Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His  
 260 265 270  
 30  
 Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe  
 275 280 285  
  
 His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn  
 290 295 300  
 35  
 Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln  
 305 310 315 320  
 40  
 Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn  
 325 330 335  
  
 Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro  
 340 345 350  
 45  
 Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala  
 355 360 365  
  
 Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly  
 370 375 380  
 50  
 Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro  
 385 390 395 400  
 55

# EP 1 310 571 A2

5 Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe  
 405 410 415  
 Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp  
 420 425 430  
 10 Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg  
 435 440 445  
 Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn Lys Asp Leu Leu Phe Ser  
 450 455 460  
 15 Arg Gly Ser Pro Ala Gly Met Ser Val Gln Pro Lys Asn Trp Leu Pro  
 465 470 475 480  
 20 Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Lys Thr Asp Asn  
 485 490 495  
 25 Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala Ser Lys Tyr Asn Leu Asn  
 500 505 510  
 Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His Lys  
 515 520 525  
 30 Asp Asp Lys Asp Lys Phe Phe Pro Met Ser Gly Val Met Ile Phe Gly  
 530 535 540  
 35 Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala Leu Asp Asn Val Met Ile  
 545 550 555 560  
 Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Arg  
 565 570 575  
 40 Phe Gly Thr Val Ala Val Asn Leu Gln Ser Ser Ser Thr Asp Pro Ala  
 580 585 590  
 45 Thr Gly Asp Val His Val Met Gly Ala Leu Pro Gly Met Val Trp Gln  
 595 600 605  
 Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His  
 610 615 620  
 50 Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu  
 625 630 635 640  
 Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala  
 645 650 655  
 55

# EP 1 310 571 A2

5 Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys Phe Ala Ser Phe Ile Thr  
 660 665 670  
  
 Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln  
 675 680 685  
 10  
 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Val Gln Tyr Thr Ser Asn  
 690 695 700  
  
 Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr Val Asp Asn Asn Gly Leu  
 705 710 715 720  
  
 Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Pro Leu  
 725 730 735  
 20  
 <210> 66  
 <211> 735  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone A3.3  
 25  
 <400> 66  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser  
 1 5 10 15  
  
 Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro  
 20 25 30  
  
 Lys Pro Asn Gln Gln His Arg Asp Asp Ser Arg Gly Leu Val Leu Pro  
 35 40 45  
 35  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
  
 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
  
 His Gln Leu Lys Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 85 90 95  
 45  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 50  
 Leu Gly Leu Val Glu Glu Ala Val Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 55

# EP 1 310 571 A2

5 Pro Ile Glu Gln Ser Pro Ala Glu Pro Asp Ser Ser Ser Gly Ile Gly  
 145 150 155 160  
 Lys Ser Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
 165 170 175  
 10 Gly Asp Thr Glu Ser Val Pro Gly Pro Gln Pro Ile Gly Glu Pro Pro  
 180 185 190  
 15 Ala Ala Pro Ser Gly Val Gly Ser Asn Thr Met Ala Ser Gly Gly Gly  
 195 200 205  
 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser  
 210 215 220  
 20 Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile  
 225 230 235 240  
 25 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
 245 250 255  
 Tyr Lys Gln Ile Ser Ser Glu Ser Gly Ala Thr Asn Asp Asn His Tyr  
 260 265 270  
 30 Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His  
 275 280 285  
 35 Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp  
 290 295 300  
 Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln Val  
 305 310 315 320  
 40 Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu  
 325 330 335  
 45 Thr Ser Ala Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr  
 340 345 350  
 Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp  
 355 360 365  
 50 Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser  
 370 375 380  
 55 Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser  
 385 390 395 400

EP 1 310 571 A2

5 Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu  
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg  
420 425 430

10 Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Lys Thr  
435 440 445

15 Gln Gly Thr Ser Gly Thr Thr Gln Gln Ser Arg Leu Gln Phe Ser Gln  
450 455 460

Ala Gly Pro Ser Ser Met Ala Gln Gln Ala Lys Asn Trp Leu Pro Gly  
465 470 475 480

20 Pro Ser Tyr Arg Gln Gln Arg Met Ser Lys Thr Ala Asn Asp Asn Asn  
485 490 495

25 Asn Ser Glu Phe Ala Trp Thr Ala Ala Thr Lys Tyr Tyr Leu Asn Gly  
500 505 510

Arg Asn Ser Leu Val Asn Pro Gly Pro Pro Val Ala Ser His Lys Asp  
515 520 525

30 Asp Glu Glu Lys Tyr Phe Pro Met His Gly Asn Leu Ile Phe Gly Lys  
530 535 540

35 Gln Gly Thr Gly Thr Thr Asn Val Asp Ile Glu Ser Val Leu Ile Thr  
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr  
565 570 575

40 Gly Gln Val Ala Thr Asn His Gln Ser Gln Asn Thr Thr Ala Ser Tyr  
580 585 590

45 Gly Ser Val Asp Ser Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp  
595 600 605

Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Thr Pro His Thr  
610 615 620

50 Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys  
625 630 635 640

55 His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn  
645 650 655

# EP 1 310 571 A2

5 Pro Ala Thr Thr Phe Thr Pro Gly Lys Phe Ala Ser Phe Ile Thr Gln  
 660 665 670  
 Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys  
 675 680 685  
 10 Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr  
 690 695 700  
 15 Asn Lys Ser Val Asn Val Glu Phe Thr Val Asp Ala Asn Gly Val Tyr  
 705 710 715 720  
 Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
 725 730 735  
 20 <210> 67  
 <211> 735  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone A3.7  
 25 <400> 57  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser  
 1 5 10 15  
 Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro  
 20 25 30  
 30 Lys Pro Asn Gln Gln His Arg Asp Asp Ser Arg Gly Leu Val Leu Pro  
 35 35 40 45  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 40 His Gln Leu Lys Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 85 90 95  
 45 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 50 Leu Gly Leu Val Glu Glu Ala Val Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 55 Pro Ile Glu Gln Ser Pro Ala Glu Pro Asp Ser Ser Ser Gly Ile Gly  
 145 150 155 160

# EP 1 310 571 A2

5 Lys Ser Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
 165 170 175  
 Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro  
 180 185 190  
 10 Ala Ala Pro Ser Gly Val Gly Ser Asn Thr Met Ala Ser Gly Gly Gly  
 195 200 205  
 15 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser  
 210 215 220  
 Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile  
 225 230 235 240  
 20 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn Arg Leu  
 245 250 255  
 25 Tyr Lys Gln Ile Ser Ser Glu Ser Gly Ala Thr Asn Asp Asn His Tyr  
 260 265 270  
 Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His  
 275 280 285  
 30 Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp  
 290 295 300  
 35 Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln Val  
 305 310 315 320  
 Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu  
 325 330 335  
 40 Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr  
 340 345 350  
 45 Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp  
 355 360 365  
 Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser  
 370 375 380  
 50 Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser  
 385 390 395 400  
 55 Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu  
 405 410 415

# EP 1 310 571 A2

5 Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg  
 420 425 430  
 Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Lys Thr  
 435 440 445  
 10 Gln Gly Thr Ser Gly Thr Thr Gln Gln Ser Arg Leu Gln Phe Ser Gln  
 450 455 460  
 15 Ala Gly Pro Ser Ser Met Ala Gln Gln Ala Lys Asn Trp Leu Pro Gly  
 465 470 475 480  
 Pro Ser Tyr Arg Gln Gln Arg Met Ser Lys Thr Ala Asn Asp Asn Asn  
 485 490 495  
 20 Asn Ser Glu Phe Ala Trp Thr Ala Ala Thr Lys Tyr Tyr Leu Asn Gly  
 500 505 510  
 25 Arg Asn Ser Leu Val Asn Pro Gly Pro Pro Met Ala Ser His Lys Asp  
 515 520 525  
 Asp Glu Glu Lys Tyr Phe Pro Met His Gly Asn Leu Ile Phe Gly Lys  
 530 535 540  
 30 Gln Gly Thr Gly Thr Thr Asn Val Asp Ile Glu Ser Val Leu Ile Thr  
 545 550 555 560  
 35 Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr  
 565 570 575  
 Gly Gln Val Ala Thr Asn His Gln Ser Gln Asn Thr Thr Ala Ser Tyr  
 580 585 590  
 40 Gly Ser Val Asp Ser Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp  
 595 600 605  
 45 Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Thr Pro His Thr  
 610 615 620  
 Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys  
 625 630 635 640  
 50 His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn  
 645 650 655  
 55 Pro Ala Thr Thr Phe Thr Pro Gly Lys Phe Ala Ser Phe Ile Thr Gln  
 660 665 670

# EP 1 310 571 A2

5 Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys  
675 680 685

Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr  
690 695 700

10 Asn Lys Ser Val Asn Val Glu Phe Thr Val Asp Ala Asn Gly Val Tyr  
705 710 715 720

15 Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
725 730 735

<210> 68  
<211> 735  
<212> PRT  
20 <213> capsid protein of AAV serotype, clone A3.4

<400> 68  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser  
1 5 10 15

25 Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro  
20 25 30

Lys Pro Asn Gln Gln His Arg Asp Asp Ser Arg Gly Leu Val Leu Pro  
35 40 45

30 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

35 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

His Gln Leu Lys Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
85 90 95

40 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
100 105 110

45 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

Leu Gly Leu Val Glu Glu Ala Val Lys Thr Ala Pro Gly Lys Lys Arg  
130 135 140

50 Pro Ile Glu Gln Ser Pro Ala Glu Pro Asp Ser Ser Ser Gly Ile Gly  
145 150 155 160

55 Glu Ser Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
165 170 175

EP 1 310 571 A2

5 Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro  
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Asn Thr Met Ala Ser Gly Gly Gly  
195 200 205

10 Ala Pro Met Ala Asp Asp Asn Glu Gly Ala Asp Gly Val Gly Asn Ser  
210 215 220

15 Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile  
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
245 250 255

20 Tyr Lys Gln Ile Ser Ser Glu Ser Gly Ala Thr Asn Asp Asn His Tyr  
260 265 270

25 Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His  
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp  
290 295 300

30 Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln Val  
305 310 315 320

35 Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu  
325 330 335

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr  
340 345 350

40 Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp  
355 360 365

45 Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser  
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser  
385 390 395 400

50 Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu  
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg  
420 425 430

55

EP 1 310 571 A2

5 Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Lys Thr  
435 440 445

Gln Gly Thr Ser Gly Thr Thr Gln Gln Ser Arg Leu Gln Phe Ser Gln  
450 455 460

10 Ala Gly Pro Ser Ser Met Ala Gln Gln Ala Lys Asn Trp Leu Pro Gly  
465 470 475 480

15 Pro Ser Tyr Arg Gln Gln Arg Met Ser Lys Thr Ala Asn Asp Asn Asn  
485 490 495

Asn Ser Glu Phe Ala Trp Thr Ala Ala Thr Lys Tyr Tyr Leu Asn Gly  
500 505 510

20 Arg Asn Ser Leu Val Asn Pro Gly Pro Pro Met Ala Ser His Lys Asp  
515 520 525

25 Asp Glu Glu Lys Tyr Phe Pro Met His Gly Asn Leu Ile Phe Gly Lys  
530 535 540

Gln Gly Thr Gly Thr Thr Asn Val Asp Ile Glu Ser Val Leu Ile Thr  
545 550 555 560

30 Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr  
565 570 575

Gly Gln Val Ala Thr Asn His Gln Ser Gln Asp Thr Thr Ala Ser Tyr  
580 585 590

Gly Ser Val Asp Ser Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp  
595 600 605

40 Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Thr Pro His Thr  
610 615 620

45 Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys  
625 630 635 640

His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn  
645 650 655

50 Pro Ala Thr Thr Phe Thr Pro Gly Lys Phe Ala Ser Phe Ile Thr Gln  
660 665 670

Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys  
675 680 685

55

# EP 1 310 571 A2

5 Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr  
690 695 700

Asn Lys Ser Val Asn Val Glu Phe Thr Val Asp Ala Asn Gly Val Tyr  
705 710 715 720

10 Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
725 730 735

15 <210> 69  
<211> 735  
<212> PRT  
<213> capsid protein of AAV serotype, clone A3.5

<400> 69  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser  
1 5 10 15

20 Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro  
20 25 30

25 Lys Pro Asn Gln Gln His Arg Asp Asp Ser Arg Gly Leu Val Leu Pro  
35 40 45

30 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

35 His Gln Leu Lys Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
85 90 95

40 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

45 Leu Gly Leu Val Glu Glu Ala Val Lys Thr Ala Pro Gly Lys Lys Arg  
130 135 140

Pro Ile Glu Gln Ser Pro Ala Glu Pro Asp Ser Ser Ser Gly Ile Gly  
145 150 155 160

50 Lys Ser Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
165 170 175

55 Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro  
180 185 190

# EP 1 310 571 A2

5  
Ala Ala Pro Ser Gly Val Gly Ser Asn Thr Met Ala Ser Gly Gly Gly  
195 200 205

10  
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser  
210 215 220

15  
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile  
225 230 235 240

20  
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
245 250 255

25  
Tyr Lys Gln Ile Ser Ser Glu Ser Gly Ala Thr Asn Asp Asn His Tyr  
260 265 270

30  
Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His  
275 280 285

35  
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp  
290 295 300

40  
Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln Val  
305 310 315 320

45  
Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu  
325 330 335

50  
Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr  
340 345 350

55  
Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp  
355 360 365

60  
Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser  
370 375 380

65  
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser  
385 390 395 400

70  
Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu  
405 410 415

75  
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg  
420 425 430

80  
Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Lys Thr  
435 440 445

# EP 1 310 571 A2

5  
Gln Gly Thr Ser Gly Thr Thr Gln Gln Ser Arg Leu Gln Phe Asn Gln  
450 455 460

10  
Ala Gly Pro Ser Ser Met Ala Gln Gln Ala Lys Asn Trp Leu Pro Gly  
465 470 475 480

15  
Pro Ser Tyr Arg Gln Gln Arg Met Ser Lys Thr Ala Asn Asp Asn Asn  
485 490 495

20  
Asn Ser Glu Phe Ala Trp Thr Ala Ala Thr Lys Tyr Tyr Pro Asn Gly  
500 505 510

25  
Arg Asn Ser Leu Val Asn Pro Gly Pro Pro Met Ala Ser His Lys Asp  
515 520 525

30  
Asp Glu Glu Lys Tyr Phe Pro Met His Gly Asn Leu Ile Phe Gly Lys  
530 535 540

35  
Gln Gly Thr Gly Thr Thr Asn Val Asp Ile Glu Ser Val Leu Ile Thr  
545 550 555 560

40  
Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr  
565 570 575

45  
Gly Gln Val Ala Thr Asn Arg Gln Ser Gln Asn Thr Thr Ala Ser Tyr  
580 585 590

50  
Gly Ser Val Asp Ser Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp  
595 600 605

55  
Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Thr Pro His Thr  
610 615 620

60  
Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys  
625 630 635 640

65  
His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn  
645 650 655

70  
Pro Ala Thr Thr Phe Thr Pro Gly Lys Phe Ala Ser Phe Ile Thr Gln  
660 665 670

75  
Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys  
675 680 685

80  
Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr  
690 695 700

# EP 1 310 571 A2

5  
Asn Lys Ser Val Asn Val Glu Phe Thr Val Asp Ala Asn Gly Val Tyr  
705 710 715 720

10  
Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
725 730 735

15  
<210> 70  
<211> 735  
<212> PRT  
<213> capsid protein of AAV serotype, clone AAV2

20  
<400> 70  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser  
1 5 10 15

25  
Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro  
20 25 30

30  
Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro  
35 40 45

35  
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

40  
Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

45  
Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
85 90 95

50  
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly  
100 105 110

55  
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

60  
Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg  
130 135 140

65  
Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly  
145 150 155 160

70  
Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr  
165 170 175

75  
Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro  
180 185 190

EP 1 310 571 A2

5           Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly  
                  195                   200                   205

          Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser  
                  210                   215                   220

10       Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile  
          225                   230                   235                   240

          Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
15                           245                   250                   255

          Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr  
                  260                   265                   270

20       Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His  
                  275                   280                   285

          Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp  
25                   290                   295                   300

          Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val  
          305                   310                   315                   320

30       Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu  
                  325                   330                   335

          Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr  
35                   340                   345                   350

          Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp  
                  355                   360                   365

40       Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser  
                  370                   375                   380

          Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser  
45                   385                   390                   395                   400

          Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu  
                  405                   410                   415

50       Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg  
                  420                   425                   430

          Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr  
                  435                   440                   445

55

# EP 1 310 571 A2

5 Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln  
 450 455 460  
 Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly  
 465 470 475 480  
 10 Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn  
 485 490 495  
 Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly  
 15 500 505 510  
 Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp  
 515 520 525  
 20 Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys  
 530 535 540  
 Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr  
 25 545 550 555 560  
 Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr  
 565 570 575  
 30 Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Arg Gln Ala Ala Thr  
 580 585 590  
 Ala Asp Val Asn Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp  
 35 595 600 605  
 Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr  
 610 615 620  
 40 Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys  
 625 630 635 640  
 His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn  
 45 645 650 655  
 Pro Ser Thr Thr Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln  
 660 665 670  
 50 Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys  
 675 680 685  
 Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr  
 690 695 700  
 55

# EP 1 310 571 A2

5 Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr  
 705 710 715 720  
  
 Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
 725 730 735  
  
 10 <210> 71  
 <211> 736  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone AAV3  
  
 <400> 71  
 15 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
  
 Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Val Pro Gln Pro  
 20 20 25 30  
  
 Lys Ala Asn Gln Gln His Gln Asp Asn Arg Arg Gly Leu Val Leu Pro  
 35 40 45  
  
 25 Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
  
 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
  
 30 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 85 90 95  
  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 35 100 105 110  
  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Ile Leu Glu Pro  
 115 120 125  
  
 40 Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Gly  
 130 135 140  
  
 Ala Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly  
 45 145 150 155 160  
  
 Lys Ser Gly Lys Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr  
 165 170 175  
  
 50 Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro  
 180 185 190  
  
 Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly  
 195 200 205  
 55

# EP 1 310 571 A2

5 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser  
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile  
225 230 235 240

10 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
245 250 255

15 Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr  
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His  
275 280 285

20 Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp  
290 295 300

25 Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val  
305 310 315 320

Arg Gly Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu  
325 330 335

30 Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr  
340 345 350

35 Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp  
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser  
370 375 380

40 Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser  
385 390 395 400

45 Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe Glu  
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg  
420 425 430

50 Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg Thr  
435 440 445

Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe Ser  
450 455 460

55

# EP 1 310 571 A2

5	Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro	465	470	475	480
	Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp Asn		485	490	495
10	Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu Asn		500	505	510
	Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys		515	520	525
15	Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe Gly		530	535	540
20	Lys Glu Gly Thr Thr Ala Ser Asn Ala Glu Leu Asp Asn Val Met Ile		545	550	555
	Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln		565	570	575
25	Tyr Gly Thr Val Ala Asn Asn Leu Gln Ser Ser Asn Thr Ala Pro Thr		580	585	590
30	Thr Gly Thr Val Asn His Gln Gly Ala Leu Pro Gly Met Val Trp Gln		595	600	605
	Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His		610	615	620
35	Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu		625	630	635
40	Lys His Pro Pro Pro Gln Ile Met Ile Lys Asn Thr Pro Val Pro Ala		645	650	655
	Asn Pro Pro Thr Thr Phe Ser Pro Ala Lys Phe Ala Ser Phe Ile Thr		660	665	670
45	Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln		675	680	685
50	Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn		690	695	700
	Tyr Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val		705	710	715
55					720

# EP 1 310 571 A2

5 Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
725 730 735

<210> 72  
<211> 737  
10 <212> PRT  
<213> capsid protein of AAV serotype, clone 3.3bVP1  
<400> 72

15 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asn Gly Arg Gly Leu Val Leu Pro  
25 35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
25 50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

30 Gln Gln Leu Asn Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
35 100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

40 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Ala Lys Lys Arg  
130 135 140

Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
45 145 150 155 160

Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln  
165 170 175

50 Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro  
180 185 190

Pro Ala Ala Pro Ser Ser Val Gly Ser Gly Thr Val Ala Ala Gly Gly  
195 200 205

55

# EP 1 310 571 A2

5 Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn  
210 215 220

Ala Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val  
225 230 235 240

10 Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His  
245 250 255

15 Leu Tyr Glu Gln Ile Ser Ser Glu Thr Ala Gly Ser Thr Asn Asp Asn  
260 265 270

Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg  
275 280 285

20 Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn  
290 295 300

25 Asn Trp Gly Phe Arg Pro Lys Lys Leu Arg Phe Lys Leu Phe Asn Ile  
305 310 315 320

Gln Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn  
325 330 335

30 Asn Leu Thr Ser Thr Ile Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu  
340 345 350

35 Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro  
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn  
370 375 380

40 Gly Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe  
385 390 395 400

45 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Ser  
405 410 415

Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu  
420 425 430

50 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala  
435 440 445

Arg Thr Gln Ser Asp Pro Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln  
450 455 460

55

# EP 1 310 571 A2

5	Phe Tyr Gln Gly Gly Pro Ser Thr Met Ala Glu Gln Ala Lys Asn Trp 465 470 475 480
	Leu Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp 485 490 495
10	Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His 500 505 510
15	Leu Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr 515 520 525
	His Lys Asp Asp Glu Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile 530 535 540
20	Phe Gly Lys Thr Gly Ala Thr Asn Lys Thr Thr Leu Glu Asn Val Leu 545 550 555 560
25	Met Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu 565 570 575
	Glu Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Asn Thr Ala Ala 580 585 590
30	Gln Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp 595 600 605
35	Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro 610 615 620
	His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly 625 630 635 640
40	Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro 645 650 655
45	Ala Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Phe Ala Ser Phe Ile 660 665 670
	Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu 675 680 685
50	Gln Lys Glu Asn Ser Lys Arg Trp Asp Pro Glu Ile Gln Tyr Thr Ser 690 695 700
55	Asn Phe Glu Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly 705 710 715 720

# EP 1 310 571 A2

5 Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn  
 725 730 735  
  
 Leu  
  
 10 <210> 73  
 <211> 644  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 223-4  
  
 <400> 73  
 15 Lys Ala Tyr Asp Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg  
 1 5 10 15  
  
 Tyr Asn His Ala Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr  
 20 20 25 30  
  
 Ser Phe Gly Gly Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg  
 35 40 45  
  
 25 Val Leu Glu Pro Leu Gly Leu Val Glu Thr Pro Ala Lys Thr Ala Pro  
 50 55 60  
  
 Gly Lys Lys Arg Pro Val Asp Ser Pro Asp Ser Thr Ser Gly Ile Gly  
 65 70 75 80  
  
 30 Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
 85 90 95  
  
 Gly Asp Ser Glu Pro Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro  
 35 100 105 110  
  
 Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly  
 115 120 125  
  
 40 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala  
 130 135 140  
  
 Ser Gly Asn Trp His Cys Asp Ser Thr Arg Leu Gly Asp Arg Val Ile  
 45 145 150 155 160  
  
 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
 165 170 175  
  
 50 Tyr Lys Gln Ile Ser Ser Gln Ser Ala Gly Ser Thr Asn Asp Asn Val  
 180 185 190  
  
 Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe  
 55 195 200 205

EP 1 310 571 A2

5 His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn  
210 215 220

Trp Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln  
225 230 235 240

10 Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn  
245 250 255

15 Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro  
260 265 270

Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala  
275 280 285

20 Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly  
290 295 300

25 Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro  
305 310 315 320

Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe  
325 330 335

30 Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Gly  
340 345 350

35 Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg  
355 360 365

Thr Gln Ser Asn Ala Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln Phe  
370 375 380

40 Tyr Gln Gly Gly Pro Thr Thr Met Ala Glu Gln Ala Lys Asn Trp Leu  
385 390 395 400

45 Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp Gln  
405 410 415

Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu  
420 425 430

50 Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His  
435 440 445

Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe  
450 455 460

55

# EP 1 310 571 A2

5 Gly Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu Met  
465 470 475 480

Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu Glu  
485 490 495

10 Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Ser Thr Ala Ala Gln  
500 505 510

15 Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp Gln  
515 520 525

Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His  
530 535 540

20 Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu  
545 550 555 560

25 Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala  
565 570 575

Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Phe Ala Ser Phe Ile Thr  
580 585 590

30 Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln  
595 600 605

35 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn  
610 615 620

Phe Asp Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly Val  
625 630 635 640

40 Tyr Ser Glu Pro

<210> 74  
<211> 644  
<212> PRT  
45 <213> capsid protein of AAV serotype, clone 223.5

<400> 74  
Lys Ala Tyr Asp Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg  
1 5 10 15

50 Tyr Asn His Ala Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr  
20 25 30

55 Ser Phe Gly Gly Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg  
35 40 45

# EP 1 310 571 A2

5 Val Leu Glu Pro Leu Gly Leu Val Glu Thr Pro Ala Lys Thr Ala Pro  
50 55 60

10 Gly Lys Lys Arg Pro Val Asp Ser Pro Asp Ser Thr Ser Gly Ile Gly  
65 70 75 80

Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
85 90 95

15 Gly Asp Ser Glu Pro Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro  
100 105 110

20 Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly  
115 120 125

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala  
130 135 140

25 Ser Gly Asn Trp His Cys Asp Ser Thr Arg Leu Gly Asp Arg Val Ile  
145 150 155 160

30 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
165 170 175

Tyr Lys Gln Ile Ser Ser Gln Ser Ala Gly Ser Thr Asn Asp Asn Val  
180 185 190

35 Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe  
195 200 205

40 His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn  
210 215 220

Trp Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln  
225 230 235 240

45 Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn  
245 250 255

50 Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro  
260 265 270

Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala  
275 280 285

55 Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly  
290 295 300

# EP 1 310 571 A2

5 Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro  
305 310 315 320

10 Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe  
325 330 335

15 Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Gly  
340 345 350

20 Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg  
355 360 365

25 Thr Gln Ser Asn Ala Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln Phe  
370 375 380

30 Tyr Gln Gly Gly Pro Thr Thr Met Ala Glu Gln Ala Lys Asn Trp Leu  
385 390 395 400

35 Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp Gln  
405 410 415

40 Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu  
420 425 430

45 Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His  
435 440 445

50 Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe  
450 455 460

55 Gly Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu Met  
465 470 475 480

60 Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu Glu  
485 490 495

65 Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Ser Thr Ala Ala Gln  
500 505 510

70 Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp Gln  
515 520 525

75 Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His  
530 535 540

80 Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu  
545 550 555 560

# EP 1 310 571 A2

5  
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala  
565 570 575

10  
Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Phe Ala Ser Phe Ile Thr  
580 585 590

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln  
595 600 605

15  
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn  
610 615 620

20  
Phe Asp Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly Val  
625 630 635 640

Tyr Ser Glu Pro

25  
<210> 75  
<211> 644  
<212> PRT  
<213> capsid protein of AAV serotype, clone 223.10

30  
<220>  
<221> MISC FEATURE  
<222> (434)..(434)  
<223> can be any amino acid

<400> 75

35  
Lys Ala Tyr Asp Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg  
1 5 10 15

Tyr Asn His Ala Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr  
20 25 30

40  
Ser Phe Gly Gly Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg  
35 40 45

45  
Val Leu Glu Pro Leu Gly Leu Val Glu Thr Pro Ala Lys Thr Ala Pro  
50 55 60

Gly Lys Lys Arg Pro Val Asp Ser Pro Asp Ser Thr Ser Gly Ile Gly  
65 70 75 80

50  
Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
85 90 95

55  
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro  
100 105 110

# EP 1 310 571 A2

5 Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly  
115 120 125

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala  
130 135 140

10 Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile  
145 150 155 160

15 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
165 170 175

Tyr Lys Gln Ile Ser Ser Gln Ser Ala Gly Ser Thr Asn Asp Asn Val  
180 185 190

20 Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe  
195 200 205

25 His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn  
210 215 220

Trp Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln  
225 230 235 240

30 Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn  
245 250 255

35 Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro  
260 265 270

Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala  
275 280 285

40 Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly  
290 295 300

45 Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro  
305 310 315 320

Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe  
325 330 335

50 Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp  
340 345 350

Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg  
355 360 365

55

# EP 1 310 571 A2

5 Thr Gln Ser Asn Ala Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln Phe  
 370 375 380  
 Tyr Gln Gly Gly Pro Thr Thr Met Ala Glu Gln Ala Lys Asn Trp Leu  
 385 390 395 400  
 10 Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp Gln  
 405 410 415  
 15 Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu  
 420 425 430  
 Asn Xaa Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His  
 435 440 445  
 20 Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe  
 450 455 460  
 25 Gly Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu Met  
 465 470 475 480  
 Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu Glu  
 485 490 495  
 30 Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Ser Thr Ala Ala Gln  
 500 505 510  
 35 Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp Gln  
 515 520 525  
 Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His  
 530 535 540  
 40 Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu  
 545 550 555 560  
 45 Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala  
 565 570 575  
 Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Phe Ala Ser Phe Ile Thr  
 580 585 590  
 50 Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln  
 595 600 605  
 55 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn  
 610 615 620

# EP 1 310 571 A2

5 Phe Asp Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly Val  
 625 630 635 640  
  
 Tyr Ser Glu Pro  
 10  
 <210> 76  
 <211> 644  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 223.2  
 15  
 <400> 76  
 Lys Ala Tyr Asp Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg  
 1 5 10 15  
  
 Tyr Asn His Ala Asp Ala Glu Phe Gln Glu Cys Leu Gln Glu Asp Thr  
 20 20 25 30  
  
 Ser Phe Gly Gly Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg  
 35 40 45  
 25  
 Val Leu Glu Pro Leu Gly Leu Val Glu Thr Pro Ala Lys Thr Ala Pro  
 50 55 60  
  
 Gly Lys Lys Arg Pro Val Asp Ser Pro Asp Ser Thr Ser Gly Ile Gly  
 65 70 75 80  
 30  
 Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
 85 90 95  
 35  
 Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro  
 100 105 110  
  
 Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Val Ala Gly Gly Gly  
 115 120 125  
 40  
 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala  
 130 135 140  
  
 Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile  
 145 150 155 160  
 45  
 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
 165 170 175  
 50  
 Tyr Lys Gln Ile Ser Ser Gln Ser Ala Gly Ser Thr Asn Asp Asn Val  
 180 185 190  
  
 Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe  
 195 200 205  
 55

# EP 1 310 571 A2

5 His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn  
210 215 220

10 Trp Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln  
225 230 235 240

Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn  
245 250 255

15 Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro  
260 265 270

20 Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala  
275 280 285

Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly  
290 295 300

25 Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro  
305 310 315 320

30 Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe  
325 330 335

Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp  
340 345 350

35 Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg  
355 360 365

40 Thr Gln Ser Asn Ala Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln Phe  
370 375 380

Tyr Gln Gly Gly Pro Thr Thr Met Ala Glu Gln Ala Lys Asn Trp Leu  
385 390 395 400

45 Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp Gln  
405 410 415

Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu  
420 425 430

50 Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His  
435 440 445

55 Lys Asp Asp Glu Glu Arg Phe Ser Pro Ser Ser Gly Val Leu Ile Phe  
450 455 460

# EP 1 310 571 A2

5 Gly Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu Met  
465 470 475 480

10 Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu Glu  
485 490 495

Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Ser Thr Ala Ala Gln  
500 505 510

15 Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp Gln  
515 520 525

20 Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His  
530 535 540

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu  
545 550 555 560

25 Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala  
565 570 575

30 Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Phe Ala Ser Phe Ile Thr  
580 585 590

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln  
595 600 605

35 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn  
610 615 620

Phe Asp Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly Val  
625 630 635 640

40 Tyr Ser Glu Pro

45 <210> 77  
<211> 644  
<212> PRT  
<213> capsid protein of AAV serotype, clone 223.7  
<400> 77

50 Lys Ala Tyr Asp Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg  
1 5 10 15

Tyr Asn His Ala Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr  
20 25 30

55

# EP 1 310 571 A2

5 Ser Phe Gly Gly Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg  
 35 40 45  
 Val Leu Glu Pro Leu Gly Leu Val Glu Thr Pro Ala Lys Thr Ala Pro  
 50 55 60  
 10 Gly Lys Lys Arg Pro Val Asp Ser Pro Asp Ser Thr Ser Gly Ile Gly  
 65 70 75 80  
 Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
 85 90 95  
 15 Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro  
 100 105 110  
 20 Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly  
 115 120 125  
 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala  
 130 135 140  
 25 Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile  
 145 150 155 160  
 30 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
 165 170 175  
 Tyr Lys Gln Ile Ser Ser Gln Ser Ala Gly Ser Thr Asn Asp Asn Val  
 180 185 190  
 35 Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe  
 195 200 205  
 40 His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn  
 210 215 220  
 Trp Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln  
 225 230 235 240  
 45 Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn  
 245 250 255  
 50 Leu Thr Ser Thr Val Gln Val Phe Ser Asp Pro Glu Tyr Gln Leu Pro  
 260 265 270  
 Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala  
 275 280 285  
 55

EP 1 310 571 A2

5 Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly  
290 295 300

Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro  
305 310 315 320

10 Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe  
325 330 335

15 Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp  
340 345 350

Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg  
355 360 365

20 Thr Gln Ser Asn Ala Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln Phe  
370 375 380

25 Tyr Gln Gly Gly Pro Thr Thr Met Ala Glu Gln Ala Lys Asn Trp Leu  
385 390 395 400

Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp Gln  
405 410 415

30 Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu  
420 425 430

35 Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His  
435 440 445

Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe  
450 455 460

40 Gly Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu Met  
465 470 475 480

45 Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu Glu  
485 490 495

Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Ser Thr Ala Ala Gln  
500 505 510

50 Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp Gln  
515 520 525

55 Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His  
530 535 540

# EP 1 310 571 A2

5 Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu  
 545 550 555 560  
  
 Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala  
 565 570 575  
 10  
 Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Ile Ala Ser Phe Ile Thr  
 580 585 590  
  
 Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln  
 595 600 605  
 15  
 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn  
 610 615 620  
  
 Phe Asp Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly Val  
 625 630 635 640  
  
 Tyr Ser Glu Pro  
 25  
 <210> 78  
 <211> 644  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 223.6  
 30  
 <400> 78  
  
 Lys Ala Tyr Asp Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg  
 1 5 10 15  
 35  
 Tyr Asn His Ala Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr  
 20 25 30  
  
 Ser Phe Gly Gly Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg  
 35 40 45  
 40  
 Val Leu Glu Pro Leu Gly Leu Val Glu Thr Pro Ala Lys Thr Ala Pro  
 50 55 60  
  
 Gly Lys Lys Arg Pro Val Asp Ser Pro Asp Ser Thr Ser Gly Ile Gly  
 65 70 75 80  
 45  
 Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
 85 90 95  
 50  
 Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro  
 100 105 110  
  
 Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly  
 115 120 125  
 55

# EP 1 310 571 A2

5 Ala Pro Met Ala Asp Asn Ser Glu Gly Ala Asp Gly Val Gly Asn Ala  
130 135 140

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile  
145 150 155 160

10 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
165 170 175

15 Tyr Lys Gln Ile Ser Ser Gln Ser Ala Gly Ser Thr Asn Asp Asn Val  
180 185 190

Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe  
195 200 205

20 His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn  
210 215 220

25 Trp Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln  
225 230 235 240

Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn  
245 250 255

30 Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro  
260 265 270

35 Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala  
275 280 285

Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly  
290 295 300

40 Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro  
305 310 315 320

45 Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe  
325 330 335

Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp  
340 345 350

50 Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg  
355 360 365

55 Thr Gln Ser Asn Ala Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln Phe  
370 375 380

EP 1 310 571 A2

5	Tyr	Gln	Gly	Gly	Pro	Thr	Thr	Met	Ala	Glu	Gln	Ala	Lys	Asn	Trp	Leu	385	390	395	400
	Pro	Gly	Pro	Cys	Phe	Arg	Gln	Gln	Arg	Val	Ser	Lys	Thr	Leu	Asp	Gln		405	410	415
10	Asn	Asn	Asn	Ser	Asn	Phe	Ala	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu		420	425	430
15	Asn	Gly	Arg	Asn	Ser	Leu	Val	Asn	Pro	Gly	Val	Ala	Met	Ala	Thr	His		435	440	445
	Lys	Asp	Asp	Glu	Glu	Arg	Phe	Phe	Pro	Ser	Ser	Gly	Val	Leu	Ile	Phe		450	455	460
20	Gly	Lys	Thr	Gly	Ala	Ala	Asn	Lys	Thr	Thr	Leu	Glu	Asn	Val	Leu	Met		465	470	475
	Thr	Asn	Glu	Glu	Glu	Ile	Arg	Pro	Thr	Asn	Pro	Val	Ala	Thr	Glu	Glu		485	490	495
25	Tyr	Gly	Ile	Val	Ser	Ser	Asn	Leu	Gln	Ala	Ala	Ser	Thr	Ala	Ala	Gln		500	505	510
30	Thr	Gln	Val	Val	Asn	Asn	Gln	Gly	Ala	Leu	Pro	Gly	Met	Val	Trp	Gln		515	520	525
35	Asn	Arg	Asp	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	Pro	His		530	535	540
	Thr	Asp	Gly	Asn	Phe	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	Gly	Leu		545	550	555
40	Lys	His	Pro	Pro	Pro	Gln	Ile	Leu	Ile	Lys	Asn	Thr	Pro	Val	Pro	Ala		565	570	575
45	Asn	Pro	Pro	Glu	Val	Phe	Thr	Pro	Ala	Lys	Leu	Ala	Ser	Phe	Ile	Thr		580	585	590
	Gln	Tyr	Ser	Thr	Gly	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	Leu	Gln		595	600	605
50	Lys	Glu	Asn	Ser	Lys	Arg	Trp	Asn	Pro	Glu	Ile	Gln	Tyr	Thr	Ser	Asn		610	615	620
55	Phe	Asp	Lys	Gln	Thr	Gly	Val	Asp	Phe	Ala	Val	Asp	Ser	Gln	Gly	Val		625	630	635
																				640

# EP 1 310 571 A2

5 Tyr Ser Glu Pro

<210> 79  
 <211> 738  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 44.1

10 <400> 79

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15

15 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 20 35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60

25 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80

30 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110

35 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125

40 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140

Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
 145 150 155 160

45 Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln  
 165 170 175

50 Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro  
 180 185 190

Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly  
 195 200 205

55

# EP 1 310 571 A2

Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser  
 210 215 220  
 5  
 Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val  
 225 230 235 240  
 10  
 Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His  
 245 250 255  
 Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp  
 260 265 270  
 15  
 Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn  
 275 280 285  
 20  
 Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn  
 290 295 300  
 Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn  
 305 310 315 320  
 25  
 Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala  
 325 330 335  
 30  
 Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln  
 340 345 350  
 Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe  
 355 360 365  
 35  
 Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn  
 370 375 380  
 40  
 Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr  
 385 390 395 400  
 Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr  
 405 410 415  
 45  
 Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser  
 420 425 430  
 50  
 Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu  
 435 440 445  
 Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu  
 450 455 460  
 55

# EP 1 310 571 A2

Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp  
 465 470 475 480  
 5  
 Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser  
 485 490 495  
 10  
 Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His  
 500 505 510  
 Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr  
 515 520 525  
 15  
 His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met  
 530 535 540  
 20  
 Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val  
 545 550 555 560  
 Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr  
 565 570 575  
 25  
 Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala  
 580 585 590  
 30  
 Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val  
 595 600 605  
 Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile  
 610 615 620  
 35  
 Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe  
 625 630 635 640  
 40  
 Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val  
 645 650 655  
 Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe  
 660 665 670  
 45  
 Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu  
 675 680 685  
 50  
 Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr  
 690 695 700  
 Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Asp  
 705 710 715 720  
 55

# EP 1 310 571 A2

Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735  
 5  
 Asn Leu  
 <210> 80  
 <211> 738  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 44.5  
 <400> 80  
 10  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 30  
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 40  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
 145 150 155 160  
 Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln  
 165 170 175  
 50  
 Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro  
 180 185 190  
 Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly  
 195 200 205  
 55

# EP 1 310 571 A2

5 Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser  
 210 215 220  
 Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val  
 225 230 235 240  
 10 Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His  
 245 250 255  
 Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp  
 15 260 265 270  
 Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn  
 275 280 285  
 20 Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn  
 290 295 300  
 Asn Asn Trp Gly Phe Arg Pro Lys Arg Pro Asn Phe Lys Leu Phe Asn  
 25 305 310 315 320  
 Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala  
 325 330 335  
 30 Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln  
 340 345 350  
 Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe  
 35 355 360 365  
 Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn  
 370 375 380  
 40 Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr  
 385 390 395 400  
 Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr  
 45 405 410 415  
 Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser  
 420 425 430  
 50 Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu  
 435 440 445  
 Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu  
 55 450 455 460

# EP 1 310 571 A2

5	Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp 465 470 475 480
10	Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser 485 490 495
15	Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His 500 505 510
20	Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr 515 520 525
25	His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met 530 535 540
30	Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val 545 550 555 560
35	Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr 565 570 575
40	Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala 580 585 590
45	Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val 595 600 605
50	Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile 610 615 620
55	Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe 625 630 635 640
60	Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val 645 650 655
65	Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe 660 665 670
70	Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu 675 680 685
75	Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr 690 695 700
80	Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Asp 705 710 715 720

# EP 1 310 571 A2

Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735  
 5  
 Asn Leu  
 10  
 <210> 81  
 <211> 738  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 44.2  
 <400> 81  
 15 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 20 25 30  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 25 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 30 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95  
 35 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 40 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 45 Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
 145 150 155 160  
 Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln  
 165 170 175  
 50 Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro  
 180 185 190  
 55 Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly  
 195 200 205

# EP 1 310 571 A2

5 Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser  
 210 215 220  
 Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val  
 225 230 235 240  
 10 Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His  
 245 250 255  
 Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp  
 15 260 265 270  
 Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn  
 275 280 285  
 20 Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn  
 290 295 300  
 Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn  
 25 305 310 315 320  
 Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala  
 325 330 335  
 30 Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln  
 340 345 350  
 Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe  
 35 355 360 365  
 Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn  
 370 375 380  
 40 Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr  
 385 390 395 400  
 Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr  
 45 405 410 415  
 Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser  
 420 425 430  
 50 Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu  
 435 440 445  
 Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu  
 55 450 455 460

# EP 1 310 571 A2

	Phe	Ser	Gln	Ala	Gly	Pro	Asn	Asn	Met	Ser	Ala	Gln	Ala	Lys	Asn	Trp	
	465					470					475					480	
5																	
	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Thr	Thr	Leu	Ser	
					485					490					495		
10	Gln	Asn	Asn	Asn	Ser	Asn	Phe	Ala	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	
				500					505					510			
	Leu	Asn	Gly	Arg	Asp	Ser	Leu	Val	Asn	Pro	Gly	Val	Ala	Met	Ala	Thr	
			515					520					525				
15																	
	His	Lys	Asp	Asp	Glu	Glu	Arg	Phe	Phe	Pro	Ser	Ser	Gly	Val	Leu	Met	
	530						535					540					
20	Phe	Gly	Lys	Gln	Gly	Ala	Gly	Lys	Asp	Asn	Val	Asp	Tyr	Ser	Ser	Val	
	545					550					555					560	
	Met	Leu	Thr	Ser	Glu	Glu	Glu	Ile	Lys	Thr	Thr	Asn	Pro	Val	Ala	Thr	
					565					570					575		
25																	
	Glu	Gln	Tyr	Gly	Val	Val	Ala	Asp	Asn	Leu	Gln	Gln	Gln	Asn	Ala	Ala	
				580					585					590			
30	Pro	Ile	Val	Gly	Ala	Val	Asn	Ser	Gln	Gly	Ala	Leu	Pro	Gly	Met	Val	
			595					600					605				
	Trp	Gln	Asn	Arg	Asp	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	
	610						615					620					
35																	
	Pro	His	Thr	Asp	Gly	Asn	Phe	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	
	625					630					635					640	
40	Gly	Leu	Lys	His	Pro	Pro	Pro	Gln	Ile	Leu	Ile	Lys	Asn	Thr	Pro	Val	
					645					650					655		
	Pro	Ala	Asp	Pro	Pro	Thr	Thr	Phe	Ser	Gln	Ala	Lys	Leu	Ala	Ser	Phe	
				660					665					670			
45																	
	Ile	Thr	Gln	Tyr	Ser	Thr	Gly	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	
			675					680					685				
50	Leu	Gln	Lys	Glu	Asn	Ser	Lys	Arg	Trp	Asn	Pro	Glu	Ile	Gln	Tyr	Thr	
	690						695					700					
	Ser	Asn	Tyr	Tyr	Lys	Ser	Thr	Asn	Val	Asp	Phe	Ala	Val	Asn	Thr	Asp	
	705					710					715					720	
55																	

# EP 1 310 571 A2

Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735  
 5  
 Asn Leu  
 <210> 82  
 <211> 738  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 29.3VP1  
 <400> 82  
 10  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 20  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 25 50 55 60  
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 30  
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 35 100 105 110  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 40  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Thr Thr Gly Ile  
 45 145 150 155 160  
 Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln  
 165 170 175  
 50  
 Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro  
 180 185 190  
 55  
 Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly  
 195 200 205

EP 1 310 571 A2

5 Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser  
210 215 220

Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val  
225 230 235 240

10 Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His  
245 250 255

15 Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp  
260 265 270

Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn  
275 280 285

20 Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn  
290 295 300

25 Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn  
305 310 315 320

Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala  
325 330 335

30 Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln  
340 345 350

35 Leu Pro Tyr Val Leu Gly Ser Ala Arg Gln Gly Cys Leu Pro Pro Phe  
355 360 365

Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn  
370 375 380

40 Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr  
385 390 395 400

45 Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr  
405 410 415

Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser  
420 425 430

50 Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu  
435 440 445

55 Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu  
450 455 460

EP 1 310 571 A2

5	Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp 465 470 475 480
10	Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser 485 490 495
15	Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His 500 505 510
20	Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr 515 520 525
25	His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met 530 535 540
30	Phe Gly Lys Gln Gly Ala Gly Lys Gly Asn Val Asp Tyr Ser Ser Val 545 550 555 560
35	Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr 565 570 575
40	Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala 580 585 590
45	Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val 595 600 605
50	Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile 610 615 620
55	Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe 625 630 635 640
60	Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val 645 650 655
65	Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe 660 665 670
70	Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu 675 680 685
75	Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr 690 695 700
80	Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Asp 705 710 715 720

# EP 1 310 571 A2

5 Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735  
  
 Asn Leu  
  
 10 <210> 83  
 <211> 738  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 29.5VP1  
 <400> 83  
  
 15 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
  
 20 Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
  
 25 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
  
 30 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
  
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95  
  
 35 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
  
 40 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
  
 45 Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
 145 150 155 160  
  
 50 Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln  
 165 170 175  
  
 Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro  
 180 185 190  
  
 55

# EP 1 310 571 A2

	Pro	Ala	Gly	Pro	Ser	Gly	Leu	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly
			195					200					205			
5	Gly	Ala	Pro	Met	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser
		210					215					220				
10	Ser	Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Gly	Val
	225					230					235					240
	Ile	Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His
					245					250					255	
15	Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ser	Thr	Asn	Asp
				260					265					270		
20	Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn
			275					280					285			
	Arg	Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn
		290					295					300				
25	Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Ser	Leu	Asn	Phe	Lys	Leu	Phe	Asn
	305					310					315					320
30	Ile	Gln	Val	Lys	Glu	Val	Thr	Gln	Asn	Glu	Gly	Thr	Lys	Thr	Ile	Ala
					325					330					335	
	Asn	Asn	Leu	Thr	Ser	Thr	Ile	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln
				340					345					350		
35	Leu	Pro	Tyr	Val	Leu	Gly	Ser	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe
			355					360					365			
40	Pro	Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn
		370					375					380				
	Asn	Gly	Ser	Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr
	385					390					395					400
45	Phe	Pro	Ser	Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr
				405						410					415	
50	Gln	Phe	Glu	Asp	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser
				420					425					430		
	Leu	Asp	Arg	Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu
			435					440					445			
55																

# EP 1 310 571 A2

	Ser	Arg	Thr	Gln	Ser	Thr	Gly	Gly	Thr	Ala	Gly	Thr	Gln	Gln	Leu	Leu	
	450						455					460					
5	Phe	Ser	Gln	Ala	Gly	Pro	Asn	Asn	Met	Ser	Ala	Gln	Ala	Lys	Asn	Trp	
	465					470				475						480	
10	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Thr	Thr	Leu	Ser	
					485					490						495	
	Gln	Asn	Asp	Asn	Ser	Asn	Phe	Ala	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	
				500					505					510			
15	Leu	Asn	Gly	Arg	Asp	Ser	Leu	Val	Asn	Pro	Gly	Val	Ala	Met	Ala	Thr	
			515				520					525					
20	His	Lys	Asp	Asp	Glu	Glu	Arg	Phe	Phe	Pro	Ser	Ser	Gly	Val	Leu	Met	
		530					535					540					
	Phe	Gly	Lys	Gln	Gly	Ala	Gly	Lys	Asp	Asn	Val	Asp	Tyr	Ser	Ser	Val	
	545				550					555						560	
25	Met	Leu	Thr	Ser	Glu	Glu	Glu	Ile	Lys	Thr	Thr	Asn	Pro	Val	Ala	Thr	
				565					570						575		
30	Glu	Gln	Tyr	Gly	Val	Val	Ala	Asp	Asn	Leu	Gln	Gln	Gln	Asn	Ala	Ala	
				580					585					590			
	Pro	Ile	Val	Gly	Ala	Val	Asn	Ser	Gln	Gly	Ala	Leu	Pro	Gly	Met	Val	
			595				600					605					
35	Trp	Gln	Asn	Arg	Asp	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	
		610				615				620							
40	Pro	His	Thr	Asp	Gly	Asn	Phe	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	
		625			630					635					640		
	Gly	Leu	Lys	His	Pro	Pro	Pro	Gln	Ile	Leu	Ile	Lys	Asn	Thr	Pro	Val	
				645				650					655				
45	Pro	Ala	Asp	Pro	Pro	Thr	Thr	Phe	Ser	Gln	Ala	Lys	Leu	Ala	Ser	Phe	
				660				665					670				
50	Ile	Thr	Gln	Tyr	Ser	Thr	Gly	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	
		675				680				685							
	Leu	Gln	Lys	Glu	Asn	Ser	Lys	Arg	Trp	Asn	Pro	Glu	Ile	Gln	Tyr	Thr	
		690				695				700							
55																	

# EP 1 310 571 A2

Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Asp  
 705 710 715 720  
  
 5 Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735  
  
 Asn Leu  
  
 10  
 <210> 84  
 <211> 738  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 42.15  
  
 15  
 <400> 84  
  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
  
 20 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 25 35 40 45  
  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
  
 30 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
  
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 35 85 90 95  
  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
  
 40 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 45 130 135 140  
  
 Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
 145 150 155 160  
  
 50 Gly Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln  
 165 170 175  
  
 Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro  
 55 180 185 190

# EP 1 310 571 A2

5	Pro	Ala	Gly	Pro	Ser	Gly	Leu	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly	195	200	205
	Gly	Ala	Pro	Met	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser	210	215	220
10	Ser	Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val	225	230	235
	Ile	Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	245	250	255
15	Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ser	Thr	Asn	Asp	260	265	270
20	Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	275	280	285
25	Arg	Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	290	295	300
	Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Asn	Phe	Lys	Leu	Phe	Asn	305	310	315
30	Ile	Gln	Val	Lys	Glu	Val	Thr	Gln	Asn	Glu	Gly	Thr	Lys	Thr	Ile	Ala	325	330	335
35	Asn	Asn	Leu	Thr	Ser	Thr	Ile	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln	340	345	350
	Leu	Pro	Tyr	Val	Leu	Gly	Ser	Ala	His	Gln	Gly	Cys	Pro	Pro	Pro	Phe	355	360	365
40	Pro	Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	370	375	380
45	Asn	Gly	Ser	Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	385	390	395
	Phe	Pro	Ser	Gln	Met	Arg	Arg	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	405	410	415
50	Gln	Phe	Glu	Asp	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	420	425	430
55	Leu	Asp	Arg	Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	435	440	445

EP 1 310 571 A2

5 Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu  
 450 455 460  
 Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp  
 465 470 475 480  
 10 Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser  
 485 490 495  
 Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His  
 500 505 510  
 15 Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr  
 515 520 525  
 His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met  
 530 535 540  
 Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val  
 545 550 555 560  
 25 Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr  
 565 570 575  
 Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala  
 580 585 590  
 Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val  
 595 600 605  
 35 Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile  
 610 615 620  
 Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe  
 625 630 635 640  
 Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val  
 645 650 655  
 Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe  
 660 665 670  
 50 Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu  
 675 680 685  
 Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr  
 690 695 700  
 55

# EP 1 310 571 A2

5 Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu  
 705 710 715 720  
  
 Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735  
  
 10 Asn Leu  
  
 <210> 85  
 <211> 738  
 <212> PRT  
 15 <213> capsid protein of AAV serotype, clone 42.8  
  
 <400> 85  
  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 20 1 5 10 15  
  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
  
 25 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 30 50 55 60  
  
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
  
 35 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95  
  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 40 100 105 110  
  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
  
 45 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
  
 Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
 50 145 150 155 160  
  
 Gly Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln  
 165 170 175  
  
 55

# EP 1 310 571 A2

	Thr	Gly	Asp	Ser	Glu	Ser	Val	Pro	Asp	Pro	Gln	Pro	Ile	Gly	Glu	Pro	
				180					185					190			
5	Pro	Ala	Gly	Pro	Ser	Gly	Leu	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly	
			195					200					205				
	Gly	Ala	Pro	Met	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser	
10		210					215					220					
	Ser	Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val	
	225					230					235					240	
15	Ile	Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	
					245					250					255		
	Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ser	Thr	Asn	Asp	
20				260					265					270			
	Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	
			275					280					285				
25	Arg	Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	
	290						295					300					
	Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Asn	Phe	Lys	Leu	Phe	Asn	
30	305					310					315					320	
	Ile	Gln	Val	Lys	Glu	Val	Thr	Gln	Asn	Glu	Gly	Thr	Lys	Thr	Ile	Ala	
					325					330					335		
35	Asn	Asn	Leu	Thr	Ser	Thr	Ile	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln	
				340				345						350			
	Leu	Pro	Tyr	Val	Leu	Gly	Ser	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	
40			355					360					365				
	Pro	Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	
		370					375					380					
45	Asn	Gly	Ser	Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	
	385					390					395					400	
	Phe	Pro	Ser	Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	
50					405					410					415		
	Gln	Phe	Glu	Asp	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	
				420					425					430			
55																	

# EP 1 310 571 A2

	Leu	Asp	Arg	Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu
			435					440					445			
5	Ser	Arg	Thr	Gln	Ser	Thr	Gly	Gly	Thr	Ala	Gly	Thr	Gln	Gln	Leu	Leu
			450				455					460				
10	Phe	Ser	Gln	Ala	Gly	Pro	Asn	Asn	Met	Ser	Ala	Gln	Ala	Lys	Asn	Trp
	465					470					475					480
	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Thr	Thr	Leu	Ser
					485					490					495	
15	Gln	Asn	Asn	Asn	Ser	Asn	Phe	Ala	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His
				500					505					510		
20	Leu	Asn	Gly	Arg	Asp	Ser	Leu	Val	Asn	Pro	Gly	Val	Ala	Met	Ala	Thr
			515					520					525			
	His	Lys	Asp	Asp	Glu	Glu	Arg	Phe	Phe	Pro	Ser	Ser	Gly	Val	Leu	Met
		530					535					540				
25	Phe	Gly	Lys	Gln	Gly	Ala	Gly	Lys	Asp	Asn	Val	Asp	Tyr	Ser	Ser	Val
	545					550					555					560
30	Met	Leu	Thr	Ser	Glu	Glu	Glu	Ile	Lys	Thr	Thr	Asn	Pro	Val	Ala	Thr
					565					570					575	
	Glu	Gln	Tyr	Gly	Val	Val	Ala	Asp	Asn	Leu	Gln	Gln	Gln	Asn	Ala	Ala
				580					585					590		
35	Pro	Ile	Val	Gly	Ala	Val	Asn	Ser	Gln	Gly	Ala	Leu	Pro	Gly	Met	Val
			595					600					605			
40	Trp	Gln	Asn	Arg	Asp	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile
		610					615					620				
	Pro	His	Thr	Asp	Gly	Asn	Phe	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe
	625					630					635					640
45	Gly	Leu	Lys	His	Pro	Pro	Pro	Gln	Ile	Leu	Ile	Lys	Asn	Thr	Pro	Val
					645					650					655	
50	Pro	Ala	Asp	Pro	Pro	Thr	Thr	Phe	Ser	Gln	Ala	Lys	Leu	Ala	Ser	Phe
				660					665					670		
	Ile	Thr	Gln	Tyr	Ser	Thr	Gly	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu
			675					680					685			
55																

# EP 1 310 571 A2

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr  
 690 695 700  
 5 Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu  
 705 710 715 720  
 Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 10 725 730 735  
 Asn Leu  
 15 <210> 86  
 <211> 733  
 <212> PRT  
 <213> amino acid of AAV serotype, clone 42.13  
 20 <400> 86  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 25 20 25 30  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 35 40 45  
 30 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 35 65 70 75 80  
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 40 85 90 95  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 45 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 50 Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln  
 145 150 155 160  
 55 Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu  
 165 170 175

# EP 1 310 571 A2

5 Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Gly Pro Ser  
 180 185 190  
 Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala  
 195 200 205  
 10 Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp  
 210 215 220  
 His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr  
 225 230 235 240  
 15 Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile  
 245 250 255  
 20 Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn Thr Tyr Phe Gly  
 260 265 270  
 Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His  
 275 280 285  
 25 Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe  
 290 295 300  
 30 Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu  
 305 310 315 320  
 Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala Asn Asn Leu Thr Ser  
 325 330 335  
 35 Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu  
 340 345 350  
 40 Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe  
 355 360 365  
 Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ala  
 370 375 380  
 45 Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met  
 385 390 395 400  
 50 Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Gln Phe Glu Asp Val  
 405 410 415  
 Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met  
 420 425 430  
 55

# EP 1 310 571 A2

	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ser	Arg	Thr	Gln	Ser	
			435					440					445				
5																	
	Thr	Gly	Gly	Thr	Ala	Gly	Thr	Gln	Gln	Leu	Leu	Phe	Ser	Gln	Ala	Gly	
		450					455					460					
10																	
	Pro	Asn	Asn	Met	Ser	Ala	Gln	Ala	Lys	Asn	Trp	Leu	Pro	Gly	Pro	Cys	
	465					470					475					480	
	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Thr	Thr	Val	Ser	Gln	Asn	Asn	Asn	Ser	
					485					490					495		
15																	
	Asn	Phe	Ala	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly	Arg	Asp	
				500					505					510			
20																	
	Ser	Leu	Val	Asn	Pro	Gly	Val	Ala	Met	Ala	Thr	His	Lys	Gly	Asp	Glu	
			515					520					525				
	Glu	Arg	Phe	Phe	Pro	Ser	Ser	Gly	Val	Leu	Met	Phe	Gly	Lys	Gln	Gly	
		530					535					540					
25																	
	Ala	Gly	Lys	Asp	Asn	Val	Asp	Tyr	Ser	Ser	Val	Met	Leu	Thr	Ser	Glu	
	545					550					555					560	
	Glu	Glu	Ile	Lys	Thr	Thr	Asn	Pro	Val	Ala	Thr	Glu	Gln	Tyr	Gly	Val	
					565					570					575		
30																	
	Val	Ala	Asp	Asn	Leu	Gln	Gln	Gln	Asn	Ala	Ala	Pro	Ile	Val	Gly	Ala	
				580					585					590			
35																	
	Val	Asn	Ser	Gln	Gly	Ala	Leu	Pro	Gly	Met	Val	Trp	Gln	Asn	Arg	Asp	
			595					600					605				
	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	Pro	His	Thr	Asp	Gly	
		610					615					620					
40																	
	Asn	Phe	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	Gly	Leu	Lys	His	Pro	
	625					630					635					640	
45																	
	Pro	Pro	Gln	Ile	Leu	Ile	Lys	Asn	Thr	Pro	Val	Pro	Ala	Asp	Pro	Pro	
				645						650					655		
50																	
	Thr	Thr	Phe	Ser	Gln	Ala	Lys	Leu	Ala	Ser	Phe	Ile	Thr	Gln	Tyr	Ser	
				660					665					670			
	Thr	Gly	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	Leu	Gln	Lys	Glu	Asn	
			675					680					685				
55																	

# EP 1 310 571 A2

Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys  
 690 695 700  
 5  
 Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu Gly Thr Tyr Ser Glu  
 705 710 715 720  
 10  
 Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Ser Leu  
 725 730  
 <210> 87  
 <211> 733  
 <212> PRT  
 15  
 <213> capsid protein of AAV serotype, clone 42.3A  
 <400> 87  
 Met Ala Ala Asp Gly His Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 20  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 25  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 30  
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 35  
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 40  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 45  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln  
 145 150 155 160  
 50  
 Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu  
 165 170 175  
 55  
 Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Gly Pro Ser  
 180 185 190

EP 1 310 571 A2

5 Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala  
195 200 205

Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp  
210 215 220

10 His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr  
225 230 235 240

Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile  
15 245 250 255

Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn Thr Tyr Phe Gly  
260 265 270

20 Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His  
275 280 285

Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Ser Trp Gly Phe  
25 290 295 300

Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu  
30 305 310 315 320

Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala Asn Asn Leu Thr Ser  
325 330 335

35 Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu  
340 345 350

Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe  
355 360 365

40 Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ala  
370 375 380

45 Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met  
385 390 395 400

Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Gln Phe Glu Asp Val  
405 410 415

50 Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met  
420 425 430

55 Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr Gln Ser  
435 440 445

EP 1 310 571 A2

5 Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu Phe Ser Gln Ala Gly  
450 455 460

Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp Leu Pro Gly Pro Cys  
465 470 475 480

10 Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser Gln Asn Asn Asn Ser  
485 490 495

15 Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asp  
500 505 510

Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His Lys Asp Asp Glu  
515 520 525

20 Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met Phe Gly Lys Gln Gly  
530 535 540

25 Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val Met Leu Thr Ser Glu  
545 550 555 560

Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr Gly Val  
565 570 575

30 Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala Pro Ile Val Gly Ala  
580 585 590

35 Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp  
595 600 605

40 Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly  
610 615 620

Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro  
625 630 635 640

45 Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asp Pro Pro  
645 650 655

Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe Ile Thr Gln Tyr Ser  
660 665 670

50 Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn  
675 680 685

55 Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys  
690 695 700

# EP 1 310 571 A2

5 Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu Gly Thr Tyr Ser Glu  
 705 710 715 720  
  
 Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
 725 730  
  
 10 <210> 88  
 <211> 731  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 42.4  
  
 15 <400> 88  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 20 25 30  
  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
  
 25 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
  
 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 30 65 70 75 80  
  
 Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 85 90 95  
  
 35 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 40 115 120 125  
  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
  
 45 Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln  
 145 150 155 160  
  
 Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu  
 50 165 170 175  
  
 Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Gly Pro Ser  
 180 185 190  
  
 55

EP 1 310 571 A2

	Gly	Leu	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly	Gly	Ala	Pro	Met	Ala	
			195					200					205				
5	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Asn	Ala	Ser	Gly	Asn	Trp	
		210					215					220					
	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val	Ile	Thr	Thr	Ser	Thr	
10						230					235					240	
	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	Leu	Tyr	Lys	Gln	Ile	
					245					250					255		
15	Ser	Ser	Gln	Ser	Gly	Ala	Thr	Asn	Asp	Asn	His	Phe	Phe	Gly	Tyr	Ser	
				260					265					270			
	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg	Phe	His	Cys	His	Phe	Ser	
20			275					280					285				
	Ser	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn	Asn	Trp	Gly	Phe	Arg	Pro	
		290					295					300					
25	Lys	Arg	Leu	Asn	Phe	Lys	Leu	Phe	Asn	Ile	Gln	Val	Lys	Glu	Val	Thr	
	305					310					315					320	
	Gln	Asn	Glu	Gly	Thr	Lys	Thr	Ile	Ala	Asn	Asn	Leu	Thr	Ser	Thr	Ile	
30					325					330					335		
	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Arg	Leu	Pro	Tyr	Val	Leu	Gly	Ser	
				340					345					350			
35	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	Pro	Ala	Asp	Val	Phe	Met	Ile	
			355					360					365				
	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asn	Gly	Ser	Gln	Ala	Val	Gly	
40		370					375					380					
	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe	Pro	Ser	Gln	Met	Leu	Arg	
	385					390					395					400	
45	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	Gln	Phe	Glu	Asp	Val	Pro	Phe	
					405					410					415		
	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	Leu	Asp	Arg	Leu	Met	Asn	Pro	
50				420					425					430			
	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ser	Arg	Thr	Gln	Ser	Thr	Gly	
			435					440					445				
55																	

# EP 1 310 571 A2

	Gly	Thr	Ala	Gly	Thr	Gln	Gln	Leu	Leu	Phe	Ser	Gln	Ala	Gly	Pro	Asn	
	450						455					460					
5	Asn	Met	Ser	Ala	Gln	Ala	Lys	Asn	Trp	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	
	465					470					475					480	
	Gln	Gln	Arg	Val	Ser	Thr	Thr	Leu	Ser	Gln	Asn	Asn	Asn	Ser	Asn	Phe	
10					485					490					495		
	Ala	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly	Arg	Asp	Ser	Leu	
				500					505					510			
15	Val	Asn	Pro	Gly	Val	Ala	Met	Ala	Thr	His	Lys	Asp	Asp	Glu	Glu	Arg	
			515					520					525				
	Phe	Phe	Pro	Ser	Ser	Gly	Val	Leu	Met	Phe	Gly	Lys	Gln	Gly	Ala	Gly	
20		530					535					540					
	Lys	Asp	Asn	Val	Asp	Tyr	Ser	Ser	Val	Met	Leu	Thr	Ser	Glu	Glu	Glu	
	545					550					555					560	
25	Ile	Lys	Thr	Thr	Asn	Pro	Val	Ala	Thr	Glu	Gln	Tyr	Gly	Val	Val	Ala	
					565					570					575		
	Asp	Asn	Leu	Gln	Gln	Gln	Asn	Ala	Ala	Pro	Ile	Val	Gly	Ala	Val	Asn	
30				580					585					590			
	Ser	Gln	Gly	Ala	Leu	Pro	Gly	Met	Val	Trp	Gln	Asn	Arg	Asp	Val	Tyr	
			595					600					605				
35	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	Pro	His	Thr	Asp	Gly	Asn	Phe	
	610						615					620					
	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	Gly	Leu	Lys	His	Pro	Pro	Pro	
40	625					630					635					640	
	Gln	Ile	Leu	Ile	Lys	Asn	Thr	Pro	Val	Pro	Ala	Asp	Pro	Pro	Thr	Thr	
					645					650					655		
45	Phe	Ser	Gln	Ala	Lys	Pro	Ala	Ser	Phe	Ile	Thr	Gln	Tyr	Ser	Thr	Gly	
				660					665					670			
	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	Leu	Gln	Lys	Glu	Asn	Ser	Lys	
50			675					680					685				
	Arg	Trp	Asn	Pro	Glu	Ile	Gln	Tyr	Thr	Ser	Asn	Tyr	Tyr	Lys	Ser	Thr	
	690						695					700					
55																	

# EP 1 310 571 A2

	Asn Val Asp Phe Ala Val Asn Thr Glu Gly Thr Tyr Ser Glu Pro Arg	
	705	715 720
5	Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu	
		725 730
	<210> 89	
	<211> 731	
10	<212> PRT	
	<213> capsid protein of AAV serotype, clone 42.5A	
	<400> 89	
15	Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser	
	1	5 10 15
	Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro	
		20 25 30
20	Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro	
		35 40 45
25	Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro	
		50 55 60
	Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp	
	65	70 75 80
30	Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala	
		85 90 95
35	Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly	
		100 105 110
	Asn Leu Gly Arg Ala Val Phe Arg Ala Lys Lys Arg Val Leu Glu Pro	
		115 120 125
40	Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg	
		130 135 140
45	Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln	
	145	150 155 160
	Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu	
		165 170 175
50	Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser	
		180 185 190
55	Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala	
		195 200 205

# EP 1 310 571 A2

Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp  
 210 215 220  
 5  
 His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr  
 225 230 235 240  
 10  
 Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile  
 245 250 255  
 15  
 Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser  
 260 265 270  
 Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser  
 275 280 285  
 20  
 Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Arg Gly Phe Arg Pro  
 290 295 300  
 25  
 Arg Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr  
 305 310 315 320  
 Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile  
 325 330 335  
 30  
 Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser  
 340 345 350  
 35  
 Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile  
 355 360 365  
 Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val Gly  
 370 375 380  
 40  
 Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg  
 385 390 395 400  
 45  
 Thr Gly Asn Asn Phe Glu Phe Ser Tyr Gln Phe Glu Asp Val Pro Phe  
 405 410 415  
 His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro  
 420 425 430  
 50  
 Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr Gln Ser Thr Gly  
 435 440 445  
 55  
 Gly Thr Ala Gly Thr Gln Gln Leu Leu Phe Ser Gln Ala Gly Pro Asn  
 450 455 460

# EP 1 310 571 A2

	Asn	Met	Ser	Ala	Gln	Ala	Lys	Asn	Trp	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	465	470	475	480
5	Gln	Gln	Arg	Val	Ser	Thr	Thr	Leu	Ser	Gln	Asn	Asn	Asn	Ser	Asn	Phe	485	490	495	
10	Ala	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly	Arg	Asp	Ser	Leu	500	505	510	
	Val	Asn	Pro	Gly	Val	Ala	Met	Ala	Thr	His	Lys	Asp	Asp	Glu	Glu	Arg	515	520	525	
15	Phe	Phe	Pro	Ser	Ser	Gly	Val	Leu	Met	Phe	Gly	Lys	Gln	Gly	Ala	Gly	530	535	540	
20	Lys	Asp	Asn	Val	Asp	Tyr	Ser	Ser	Val	Met	Leu	Thr	Ser	Glu	Glu	Glu	545	550	555	560
	Ile	Lys	Thr	Thr	Asn	Pro	Val	Ala	Thr	Glu	Gln	Tyr	Gly	Val	Val	Ala	565	570	575	
25	Asp	Asn	Leu	Gln	Gln	Gln	Asn	Ala	Ala	Pro	Ile	Val	Gly	Ala	Val	Asn	580	585	590	
30	Ser	Gln	Gly	Ala	Leu	Pro	Gly	Met	Ala	Trp	Gln	Asn	Arg	Asp	Val	Tyr	595	600	605	
	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	Pro	His	Thr	Asp	Gly	Asn	Phe	610	615	620	
35	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	Gly	Leu	Lys	His	Pro	Pro	Pro	625	630	635	640
40	Gln	Ile	Leu	Ile	Lys	Asn	Thr	Pro	Val	Pro	Ala	Asp	Pro	Pro	Thr	Thr	645	650	655	
	Phe	Ser	Gln	Ala	Lys	Leu	Ala	Ser	Phe	Ile	Thr	Gln	Tyr	Ser	Thr	Gly	660	665	670	
45	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	Leu	Gln	Lys	Glu	Asn	Ser	Lys	675	680	685	
50	Arg	Trp	Asn	Pro	Glu	Ile	Gln	Tyr	Thr	Ser	Asn	Tyr	Tyr	Lys	Ser	Thr	690	695	700	
55	Asn	Val	Asp	Phe	Ala	Val	Asn	Thr	Glu	Gly	Thr	Tyr	Ser	Glu	Pro	Arg	705	710	715	720

# EP 1 310 571 A2

Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
725 730

5  
<210> 90  
<211> 733  
<212> PRT  
<213> capsid protein of AAV serotype, clone 42.1B

10  
<400> 90  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

15  
Glu Gly Ile Arg Glu Trp Trp Asp Leu Arg Pro Gly Ala Pro Lys Pro  
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

20  
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

25  
Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
85 90 95

30  
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
100 105 110

35  
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
130 135 140

40  
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln  
145 150 155 160

45  
Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu  
165 170 175

Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Gly Pro Ser  
180 185 190

50  
Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala  
195 200 205

55  
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp  
210 215 220

# EP 1 310 571 A2

His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr  
 225 230 235 240  
 5  
 Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile  
 245 250 255  
 10  
 Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn Thr Tyr Phe Gly  
 260 265 270  
 Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His  
 275 280 285  
 15  
 Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe  
 290 295 300  
 20  
 Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu  
 305 310 315 320  
 Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala Asn Asn Leu Thr Ser  
 325 330 335  
 25  
 Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu  
 340 345 350  
 Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe  
 355 360 365  
 30  
 Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ala  
 370 375 380  
 35  
 Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met  
 385 390 395 400  
 Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Gln Phe Glu Asp Val  
 405 410 415  
 Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met  
 420 425 430  
 45  
 Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr Gln Ser  
 435 440 445  
 Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu Phe Ser Gln Ala Gly  
 450 455 460  
 50  
 Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp Leu Pro Gly Pro Cys  
 465 470 475 480  
 55

# EP 1 310 571 A2

	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Thr	Thr	Val	Ser	Gln	Asn	Asn	Asn	Ser	
					485					490					495		
5	Asn	Phe	Ala	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly	Arg	Asp	
				500					505					510			
	Ser	Leu	Val	Asn	Pro	Gly	Val	Ala	Met	Ala	Thr	His	Lys	Gly	Asp	Glu	
10			515					520					525				
	Glu	Arg	Phe	Phe	Pro	Ser	Ser	Gly	Val	Leu	Met	Phe	Gly	Lys	Gln	Gly	
		530					535					540					
15	Ala	Gly	Lys	Asp	Asn	Val	Asp	Tyr	Ser	Ser	Val	Met	Leu	Thr	Ser	Glu	
	545					550					555					560	
	Glu	Glu	Ile	Lys	Thr	Thr	Asn	Pro	Val	Ala	Thr	Glu	Gln	Tyr	Gly	Val	
20					565					570					575		
	Val	Ala	Asp	Asn	Leu	Gln	Gln	Gln	Asn	Ala	Ala	Pro	Ile	Val	Gly	Ala	
				580					585					590			
25	Val	Asn	Ser	Gln	Gly	Ala	Leu	Pro	Gly	Met	Val	Trp	Gln	Asn	Arg	Asp	
			595					600					605				
	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	Pro	His	Thr	Asp	Gly	
30		610					615					620					
	Asn	Phe	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	Gly	Leu	Lys	His	Pro	
	625					630					635					640	
35	Pro	Pro	Gln	Ile	Leu	Ile	Lys	Asn	Thr	Pro	Val	Pro	Ala	Asp	Pro	Pro	
					645					650					655		
	Thr	Thr	Phe	Ser	Gln	Ala	Lys	Leu	Ala	Ser	Phe	Ile	Thr	Gln	Tyr	Ser	
40				660					665					670			
	Thr	Gly	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	Leu	Gln	Lys	Glu	Asn	
			675					680					685				
45	Ser	Lys	Arg	Trp	Asn	Pro	Glu	Ile	Gln	Tyr	Thr	Ser	Asn	Tyr	Tyr	Lys	
		690					695					700					
	Ser	Thr	Asn	Val	Asp	Phe	Ala	Val	Asn	Thr	Glu	Gly	Thr	Tyr	Ser	Glu	
50						710					715					720	
	Pro	Arg	Pro	Ile	Gly	Thr	Arg	Tyr	Leu	Thr	Arg	Asn	Leu				
					725					730							
55																	

# EP 1 310 571 A2

<210> 91  
 <211> 738  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 42.5B

5  
 <400> 91  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15

10  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45

15  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80

20  
 Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 85 90 95

25  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125

30  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140

35  
 Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
 145 150 155 160

Gly Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln  
 165 170 175

40  
 Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro  
 180 185 190

45  
 Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly  
 195 200 205

Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser  
 210 215 220

50  
 Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val  
 225 230 235 240

55  
 Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His  
 245 250 255

# EP 1 310 571 A2

5	Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp	260	265	270
	Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn	275	280	285
10	Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn	290	295	300
15	Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn	305	310	315
	Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala	325	330	335
20	Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln	340	345	350
25	Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe	355	360	365
	Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn	370	375	380
30	Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr	385	390	395
35	Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr	405	410	415
	Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser	420	425	430
40	Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu	435	440	445
45	Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu	450	455	460
	Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp	465	470	475
50	Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser	485	490	495
55	Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His	500	505	510

# EP 1 310 571 A2

5 Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr  
 515 520 525  
 His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met  
 530 535 540  
 10 Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val  
 545 550 555 560  
 Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr  
 565 570 575  
 15 Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala  
 580 585 590  
 20 Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val  
 595 600 605  
 Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile  
 610 615 620  
 25 Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe  
 625 630 635 640  
 30 Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val  
 645 650 655  
 Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe  
 660 665 670  
 35 Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu  
 675 680 685  
 40 Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr  
 690 695 700  
 Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu  
 705 710 715 720  
 45 Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735  
 50 Asn Leu  
 <210> 92  
 <211> 738  
 <212> PRT  
 55 <213> capsid protein of AAV serotype, clone 43.1

# EP 1 310 571 A2

<400> 92  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

5  
Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

10  
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

15  
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

20  
Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

25  
Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
85 90 95

30  
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
100 105 110

35  
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

40  
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
130 135 140

45  
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
145 150 155 160

50  
Gly Lys Lys Gly His Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln  
165 170 175

55  
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro  
180 185 190

60  
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly  
195 200 205

65  
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser  
210 215 220

70  
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val  
225 230 235 240

75  
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His  
245 250 255

# EP 1 310 571 A2

	Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ser	Thr	Asn	Asp	
				260					265					270			
5	Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	
			275					280					285				
	Arg	Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	
		290					295					300					
10																	
	Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Asn	Phe	Lys	Leu	Phe	Asn	
	305					310					315					320	
15	Ile	Gln	Val	Lys	Glu	Val	Thr	Gln	Asn	Glu	Gly	Thr	Lys	Thr	Ile	Ala	
					325					330					335		
	Asn	Asn	Leu	Thr	Ser	Thr	Ile	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln	
				340					345					350			
20																	
	Leu	Pro	Tyr	Val	Pro	Gly	Ser	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	
			355					360					365				
25	Pro	Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	
		370					375					380					
	Asn	Gly	Ser	Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	
	385					390					395					400	
30																	
	Phe	Pro	Ser	Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	
					405					410					415		
35	Thr	Phe	Glu	Asp	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	
				420					425					430			
	Leu	Asp	Arg	Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	
			435					440					445				
40																	
	Ser	Arg	Thr	Gln	Ser	Thr	Gly	Gly	Thr	Gln	Gly	Thr	Gln	Gln	Leu	Leu	
		450					455					460					
45	Phe	Ser	Gln	Ala	Gly	Pro	Ala	Asn	Met	Ser	Ala	Gln	Ala	Lys	Asn	Trp	
	465					470					475					480	
	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Thr	Thr	Leu	Ser	
					485					490					495		
50																	
	Gln	Asn	Asn	Asn	Ser	Asn	Phe	Ala	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	
				500					505					510			
55	Leu	Asn	Gly	Arg	Asp	Ser	Leu	Val	Asn	Pro	Gly	Val	Ala	Met	Ala	Thr	
			515					520					525				

# EP 1 310 571 A2

5 His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met  
 530 535 540  
 Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val  
 545 550 555 560  
 10 Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr  
 565 570 575  
 15 Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Thr Asn Gly Ala  
 580 585 590  
 Pro Ile Val Gly Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val  
 595 600 605  
 20 Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile  
 610 615 620  
 25 Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe  
 625 630 635 640  
 Gly Leu Lys His Pro Pro Pro Gln Ile Leu Val Lys Asn Thr Pro Val  
 645 650 655  
 30 Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe  
 660 665 670  
 35 Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu  
 675 680 685  
 Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr  
 690 695 700  
 40 Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu  
 705 710 715 720  
 45 Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735  
 Asn Leu  
 50  
 <210> 93  
 <211> 738  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 43.12  
 55

EP 1 310 571 A2

<400> 93  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

5  
Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

10  
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

15  
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

20  
Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

25  
Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
85 90 95

30  
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
100 105 110

35  
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

40  
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
130 135 140

45  
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
145 150 155 160

50  
Gly Lys Lys Gly His Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln  
165 170 175

55  
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro  
180 185 190

60  
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly  
195 200 205

65  
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser  
210 215 220

70  
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val  
225 230 235 240

75  
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His  
245 250 255

# EP 1 310 571 A2

	Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ser	Thr	Asn	Asp	
				260					265					270			
5	Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	
			275					280					285				
	Arg	Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	
10		290					295					300					
	Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Asn	Phe	Lys	Leu	Phe	Asn	
	305					310					315					320	
15	Ile	Gln	Val	Lys	Glu	Val	Thr	Gln	Asn	Glu	Gly	Thr	Lys	Thr	Ile	Ala	
					325					330					335		
	Asn	Asn	Leu	Thr	Ser	Thr	Ile	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln	
20				340					345					350			
	Leu	Pro	Tyr	Val	Leu	Gly	Ser	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	
			355					360					365				
25	Pro	Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	
		370					375					380					
	Asn	Gly	Ser	Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	
30	385					390					395					400	
	Phe	Pro	Ser	Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	
					405					410					415		
35	Thr	Phe	Glu	Asp	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	
				420					425					430			
	Leu	Asp	Arg	Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	
40			435					440					445				
	Ser	Arg	Thr	Gln	Ser	Thr	Gly	Gly	Thr	Gln	Gly	Thr	Gln	Gln	Leu	Leu	
		450					455					460					
45	Phe	Ser	Gln	Ala	Gly	Pro	Ala	Asn	Met	Ser	Ala	Gln	Ala	Lys	Asn	Trp	
	465					470					475					480	
	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Thr	Thr	Leu	Ser	
50					485					490					495		
	Gln	Asn	Asn	Asn	Ser	Asn	Phe	Ala	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	
				500					505					510			
55																	

**EP 1 310 571 A2**

**Amino acid sequence**

Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr  
515 520 525

His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met  
530 535 540

Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val  
545 550 555 560

Met Leu Thr Ser Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr  
565 570 575

Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Thr Asn Gly Ala  
580 585 590

Pro Ile Val Gly Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val  
595 600 605

Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile  
610 615 620

Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe  
625 630 635 640

Gly Leu Lys His Pro Pro Pro Gln Ile Leu Val Lys Asn Thr Pro Val  
645 650 655

Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe  
660 665 670

Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu  
675 680 685

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr  
690 695 700

Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu  
705 710 715 720

Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
725 730 735

Asn Leu

<210> 94  
<211> 738  
<212> PRT  
<213> capsid protein of AAV serotype, clone 43.5

# EP 1 310 571 A2

<400> 94  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15  
5  
Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30  
10  
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45  
15  
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60  
20  
Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80  
30  
Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
85 90 95  
25  
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
100 105 110  
30  
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125  
35  
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
130 135 140  
40  
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
145 150 155 160  
45  
Gly Lys Lys Gly His Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln  
165 170 175  
50  
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro  
180 185 190  
55  
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly  
195 200 205  
60  
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser  
210 215 220  
65  
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val  
225 230 235 240  
70  
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His  
245 250 255

# EP 1 310 571 A2

Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp  
 260 265 270  
 5  
 Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn  
 275 280 285  
 10  
 Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn  
 290 295 300  
 Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn  
 305 310 315 320  
 15  
 Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala  
 325 330 335  
 20  
 Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln  
 340 345 350  
 Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe  
 355 360 365  
 25  
 Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn  
 370 375 380  
 30  
 Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr  
 385 390 395 400  
 Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr  
 405 410 415  
 35  
 Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser  
 420 425 430  
 40  
 Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu  
 435 440 445  
 Ser Arg Thr Gln Ser Thr Gly Gly Thr Gln Gly Thr Gln Gln Leu Leu  
 450 455 460  
 45  
 Phe Ser Gln Ala Gly Pro Ala Asn Met Ser Ala Gln Ala Lys Asn Trp  
 465 470 475 480  
 50  
 Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser  
 485 490 495  
 55  
 Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His  
 500 505 510

# EP 1 310 571 A2

5 Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr  
 515 520 525  
 His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met  
 530 535 540  
 10 Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val  
 545 550 555 560  
 Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr  
 565 570 575  
 15 Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Thr Asn Gly Ala  
 580 585 590  
 20 Pro Ile Val Gly Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val  
 595 600 605  
 Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile  
 610 615 620  
 25 Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe  
 625 630 635 640  
 30 Gly Leu Lys His Pro Pro Pro Gln Ile Leu Val Lys Asn Thr Pro Val  
 645 650 655  
 Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe  
 660 665 670  
 35 Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu  
 675 680 685  
 40 Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr  
 690 695 700  
 Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu  
 705 710 715 720  
 45 Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735  
 50 Asn Leu  
 <210> 95  
 <211> 738  
 <212> PRT  
 55 <213> capsid protein of AAV serotype, clone AAV8

EP 1 310 571 A2

<400> 95

5 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

10 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

15 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

20 Gln Gln Leu Gln Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
85 90 95

25 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

30 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
130 135 140

35 Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
145 150 155 160

Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln  
165 170 175

40 Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro  
180 185 190

45 Pro Ala Ala Pro Ser Gly Val Gly Pro Asn Thr Met Ala Ala Gly Gly  
195 200 205

Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser  
210 215 220

50 Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val  
225 230 235 240

55

EP 1 310 571 A2

	Ile	Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	
					245					250					255		
5	Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ala	Thr	Asn	Asp	
				260					265					270			
	Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	
10			275					280					285				
	Arg	Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	
		290					295					300					
15	Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Ser	Phe	Lys	Leu	Phe	Asn	
	305					310					315					320	
	Ile	Gln	Val	Lys	Glu	Val	Thr	Gln	Asn	Glu	Gly	Thr	Lys	Thr	Ile	Ala	
20					325					330					335		
	Asn	Asn	Leu	Thr	Ser	Thr	Ile	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln	
				340					345					350			
25	Leu	Pro	Tyr	Val	Leu	Gly	Ser	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	
			355					360					365				
	Pro	Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	
30		370					375					380					
	Asn	Gly	Ser	Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	
	385					390					395					400	
35	Phe	Pro	Ser	Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Gln	Phe	Thr	Tyr	
					405					410					415		
	Thr	Phe	Glu	Asp	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	
40				420					425					430			
	Leu	Asp	Arg	Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	
			435					440					445				
45	Ser	Arg	Thr	Gln	Thr	Thr	Gly	Gly	Thr	Ala	Asn	Thr	Gln	Thr	Leu	Gly	
		450					455					460					
	Phe	Ser	Gln	Gly	Gly	Pro	Asn	Thr	Met	Ala	Asn	Gln	Ala	Lys	Asn	Trp	
50						470					475					480	
	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Thr	Thr	Thr	Gly	
					485					490						495	
55																	

# EP 1 310 571 A2

Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Ala Gly Thr Lys Tyr His  
 500 505 510  
 5  
 Leu Asn Gly Arg Asn Ser Leu Ala Asn Pro Gly Ile Ala Met Ala Thr  
 515 520 525  
 10  
 His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Asn Gly Ile Leu Ile  
 530 535 540  
 Phe Gly Lys Gln Asn Ala Ala Arg Asp Asn Ala Asp Tyr Ser Asp Val  
 545 550 555 560  
 15  
 Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr  
 565 570 575  
 20  
 Glu Glu Tyr Gly Ile Val Ala Asp Asn Leu Gln Gln Gln Asn Thr Ala  
 580 585 590  
 Pro Gln Ile Gly Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val  
 595 600 605  
 25  
 Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile  
 610 615 620  
 30  
 Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe  
 625 630 635 640  
 Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val  
 645 650 655  
 35  
 Pro Ala Asp Pro Pro Thr Thr Phe Asn Gln Ser Lys Leu Asn Ser Phe  
 660 665 670  
 40  
 Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu  
 675 680 685  
 Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr  
 690 695 700  
 45  
 Ser Asn Tyr Tyr Lys Ser Thr Ser Val Asp Phe Ala Val Asn Thr Glu  
 705 710 715 720  
 50  
 Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg  
 725 730 735  
 Asn Leu

55

# EP 1 310 571 A2

<210> 96  
 <211> 736  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 43.21  
 5  
 <400> 96  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 10  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 15  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 20  
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 25  
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 30  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 35  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly  
 145 150 155 160  
 40  
 Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
 165 170 175  
 45  
 Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro  
 180 185 190  
 Ala Ala Pro Ser Gly Leu Gly Pro Asn Thr Met Ala Ser Gly Gly Gly  
 195 200 205  
 50  
 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser  
 210 215 220  
 55

# EP 1 310 571 A2

	Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile	
	225	230 235 240
5	Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu	
		245 250 255
10	Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn	
		260 265 270
	Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg	
		275 280 285
15	Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn	
		290 295 300
20	Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile	
		305 310 315 320
	Gln Val Lys Glu Val Thr Thr Asn Glu Gly Thr Lys Thr Ile Ala Asn	
		325 330 335
25	Asn Leu Thr Ser Thr Val Arg Val Phe Thr Asp Ser Glu Tyr Gln Leu	
		340 345 350
30	Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro	
		355 360 365
	Ala Asp Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn	
		370 375 380
35	Gly Ser Gln Ala Leu Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe	
		385 390 395 400
40	Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr	
		405 410 415
	Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu	
		420 425 430
45	Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Val	
		435 440 445
50	Arg Thr Gln Thr Thr Gly Thr Gly Gly Thr Gln Thr Leu Ala Phe Ser	
		450 455 460
	Gln Ala Gly Pro Ser Ser Met Ala Asn Gln Ala Arg Asn Trp Val Pro	
		465 470 475 480
55		

# EP 1 310 571 A2

	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Thr	Thr	Thr	Asn	Gln	Ser	
					485					490					495		
5	Asn	Asn	Ser	Asn	Phe	Ala	Trp	Thr	Gly	Ala	Ala	Lys	Phe	Lys	Leu	Asn	
				500					505					510			
	Gly	Arg	Asp	Ser	Leu	Met	Asn	Pro	Gly	Val	Ala	Met	Ala	Ser	His	Lys	
10			515					520					525				
	Asp	Asp	Asp	Asp	Arg	Phe	Phe	Pro	Ser	Ser	Gly	Val	Leu	Ile	Phe	Gly	
		530					535					540					
15	Lys	Gln	Gly	Ala	Gly	Asn	Asp	Gly	Val	Asp	Tyr	Ser	Gln	Val	Leu	Ile	
	545					550					555					560	
	Thr	Asp	Glu	Glu	Glu	Ile	Lys	Ala	Thr	Asn	Pro	Val	Ala	Thr	Glu	Glu	
20					565					570					575		
	Tyr	Gly	Ala	Val	Ala	Ile	Asn	Asn	Gln	Ala	Ala	Asn	Thr	Gln	Ala	Gln	
				580					585					590			
25	Thr	Gly	Leu	Val	His	Asn	Gln	Gly	Val	Ile	Pro	Gly	Met	Val	Trp	Gln	
			595					600					605				
	Asn	Arg	Asp	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	Pro	His	
30		610					615					620					
	Thr	Asp	Gly	Asn	Phe	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	Gly	Leu	
	625					630					635					640	
35	Lys	His	Pro	Pro	Pro	Gln	Ile	Leu	Ile	Lys	Asn	Thr	Pro	Val	Pro	Ala	
					645					650					655		
	Asp	Pro	Pro	Leu	Thr	Phe	Asn	Gln	Ala	Lys	Leu	Asn	Ser	Phe	Ile	Thr	
40				660					665					670			
	Gln	Tyr	Ser	Thr	Gly	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	Leu	Gln	
			675					680					685				
45	Lys	Glu	Asn	Ser	Lys	Arg	Trp	Asn	Pro	Glu	Ile	Gln	Tyr	Thr	Ser	Asn	
	690						695					700					
	Tyr	Tyr	Lys	Ser	Thr	Asn	Val	Asp	Phe	Ala	Val	Asn	Thr	Glu	Gly	Val	
50	705					710					715					720	
	Tyr	Ser	Glu	Pro	Arg	Pro	Ile	Gly	Thr	Arg	Tyr	Leu	Thr	Arg	Asn	Leu	
					725					730					735		
55																	

# EP 1 310 571 A2

```

<210> 97
<211> 736
<212> PRT
5 <213> capsid protein of AAV serotype, clone 43.25

<400> 97
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

10 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

15 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

20 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

25 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

30 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
35

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
145 150 155 160

40 Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
180 185 190
45

Ala Ala Pro Ser Gly Leu Gly Pro Asn Thr Met Ala Ser Gly Gly Gly
195 200 205

50 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

55 Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile
225 230 235 240

```

# EP 1 310 571 A2

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
 245 250 255  
 5  
 Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn  
 260 265 270  
 10  
 Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg  
 275 280 285  
 Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn  
 290 295 300  
 15  
 Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile  
 305 310 315 320  
 20  
 Gln Val Lys Glu Val Thr Thr Asn Glu Gly Thr Lys Thr Ile Ala Asn  
 325 330 335  
 Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu  
 340 345 350  
 25  
 Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro  
 355 360 365  
 30  
 Ala Asp Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn  
 370 375 380  
 Gly Ser Gln Ala Leu Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe  
 385 390 395 400  
 35  
 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr  
 405 410 415  
 40  
 Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu  
 420 425 430  
 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Val  
 435 440 445  
 45  
 Arg Thr Gln Thr Thr Gly Thr Gly Gly Thr Gln Thr Leu Ala Phe Ser  
 450 455 460  
 50  
 Gln Ala Gly Pro Ser Ser Met Ala Asn Gln Ala Arg Asn Trp Val Pro  
 465 470 475 480  
 Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Asn Gln Asn  
 485 490 495  
 55

# EP 1 310 571 A2

5 Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Ala Lys Phe Lys Leu Asn  
 500 505 510  
 Gly Arg Asp Ser Leu Met Asn Pro Gly Val Ala Met Ala Ser His Lys  
 515 520 525  
 10 Asp Asp Asp Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe Gly  
 530 535 540  
 Lys Gln Gly Ala Gly Asn Asp Gly Val Asp Tyr Ser Gln Val Leu Ile  
 545 550 555 560  
 15 Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Glu  
 565 570 575  
 20 Tyr Gly Ala Val Ala Ile Asn Asn Gln Ala Ala Asn Thr Gln Ala Gln  
 580 585 590  
 Thr Gly Leu Val His Asn Gln Gly Val Ile Pro Gly Met Val Trp Gln  
 595 600 605  
 25 Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His  
 610 615 620  
 30 Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu  
 625 630 635 640  
 Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala  
 645 650 655  
 35 Asp Pro Pro Leu Thr Phe Asn Gln Ala Lys Leu Asn Ser Phe Ile Thr  
 660 665 670  
 40 Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln  
 675 680 685  
 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn  
 690 695 700  
 45 Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu Gly Val  
 705 710 715 720  
 50 Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
 725 730 735  
 55

# EP 1 310 571 A2

<210> 98  
 <211> 736  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 43.23  
 5  
 <400> 98  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 10  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 15  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 20  
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 25  
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 30  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 35  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly  
 145 150 155 160  
 40  
 Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
 165 170 175  
 45  
 Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro  
 180 185 190  
 Ala Ala Pro Ser Gly Leu Gly Pro Asn Thr Met Ala Ser Gly Gly Gly  
 195 200 205  
 50  
 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser  
 210 215 220  
 55  
 Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile  
 225 230 235 240

EP 1 310 571 A2

5	Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu	245	250	255
	Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn	260	265	270
10	Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg	275	280	285
15	Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn	290	295	300
	Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile	305	310	315
20	Gln Val Lys Glu Val Thr Thr Asn Glu Gly Thr Lys Thr Ile Ala Asn	325	330	335
25	Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Leu Glu Tyr Gln Leu	340	345	350
	Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro	355	360	365
30	Ala Asp Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn	370	375	380
35	Gly Ser Gln Ala Leu Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe	385	390	395
	Pro Ser Gln Met Pro Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr	405	410	415
40	Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu	420	425	430
45	Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Val	435	440	445
	Arg Thr Gln Thr Thr Gly Thr Gly Gly Thr Gln Thr Leu Ala Phe Ser	450	455	460
50	Gln Ala Gly Pro Ser Ser Met Ala Asn Gln Ala Arg Asn Trp Val Pro	465	470	475
55	Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Asn Gln Asn	485	490	495

EP 1 310 571 A2

5 Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Ala Lys Phe Lys Leu Asn  
 500 505 510  
 Gly Arg Asp Ser Leu Met Asn Pro Gly Val Ala Met Ala Ser His Lys  
 515 520 525  
 10 Asp Asp Asp Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe Gly  
 530 535 540  
 15 Lys Gln Gly Ala Gly Asn Asp Gly Val Asp Tyr Ser Gln Val Leu Ile  
 545 550 555 560  
 Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Glu  
 565 570 575  
 20 Tyr Gly Ala Val Ala Ile Asn Asn Gln Ala Ala Asn Thr Gln Ala Gln  
 580 585 590  
 25 Thr Gly Leu Val His Asn Gln Gly Val Ile Pro Gly Met Val Trp Gln  
 595 600 605  
 Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His  
 610 615 620  
 30 Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu  
 625 630 635 640  
 35 Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala  
 645 650 655  
 Asp Pro Pro Leu Thr Phe Asn Gln Ala Lys Leu Asn Ser Phe Ile Thr  
 660 665 670  
 40 Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln  
 675 680 685  
 45 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn  
 690 695 700  
 Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu Gly Val  
 705 710 715 720  
 50 Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
 725 730 735  
 55

# EP 1 310 571 A2

<210> 99  
 <211> 736  
 <212> PRT  
 5 <213> capsid protein of AAV serotype, clone 43.20  
  
 <400> 99  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 10 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 15 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 20 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
 85 90 95  
 25 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 30 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 35 Leu Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly  
 145 150 155 160  
 40 Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
 165 170 175  
 Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro  
 180 185 190  
 45 Ala Ala Pro Ser Gly Leu Gly Pro Asn Thr Met Ala Ser Gly Gly Gly  
 195 200 205  
 50 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser  
 210 215 220  
 55 Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile  
 225 230 235 240

# EP 1 310 571 A2

5 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
 245 250 255  
 Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn  
 260 265 270  
 10 Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg  
 275 280 285  
 Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn  
 290 295 300  
 15 Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile  
 305 310 315 320  
 Gln Val Lys Glu Val Thr Thr Asn Glu Gly Thr Lys Thr Ile Ala Asn  
 325 330 335  
 20 Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu  
 340 345 350  
 25 Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro  
 355 360 365  
 30 Ala Asp Val Phe Thr Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn  
 370 375 380  
 Gly Ser Gln Ala Leu Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe  
 385 390 395 400  
 35 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr  
 405 410 415  
 40 Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu  
 420 425 430  
 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Val  
 435 440 445  
 45 Arg Thr Gln Thr Thr Gly Thr Gly Gly Thr Gln Thr Leu Ala Phe Ser  
 450 455 460  
 50 Gln Ala Gly Pro Ser Ser Met Ala Asn Gln Ala Arg Asn Trp Val Pro  
 465 470 475 480  
 Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Asn Gln Asn  
 485 490 495  
 55

# EP 1 310 571 A2

	Asn	Asn	Ser	Asn	Phe	Ala	Trp	Thr	Gly	Ala	Ala	Lys	Phe	Lys	Leu	Asn	
				500					505					510			
5	Gly	Arg	Asp	Ser	Leu	Met	Asn	Pro	Gly	Val	Ala	Met	Ala	Ser	His	Lys	
			515					520					525				
	Asp	Asp	Asp	Asp	Arg	Phe	Phe	Pro	Ser	Ser	Gly	Val	Leu	Ile	Phe	Gly	
10		530					535					540					
	Lys	Gln	Gly	Ala	Gly	Asn	Asp	Gly	Val	Asp	Tyr	Ser	Gln	Val	Leu	Ile	
	545					550					555					560	
15	Thr	Asp	Glu	Glu	Glu	Ile	Lys	Ala	Thr	Asn	Pro	Val	Ala	Thr	Glu	Glu	
					565					570					575		
	Tyr	Gly	Ala	Val	Ala	Ile	Asn	Asn	Gln	Ala	Ala	Asn	Thr	Gln	Ala	Gln	
20				580					585					590			
	Thr	Gly	Leu	Val	His	Asn	Gln	Gly	Val	Ile	Pro	Gly	Met	Val	Trp	Gln	
			595					600					605				
25	Asn	Arg	Asp	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	Pro	His	
		610					615					620					
	Thr	Asp	Gly	Asn	Phe	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	Gly	Leu	
30						630					635					640	
	Lys	His	Pro	Pro	Pro	Gln	Ile	Leu	Ile	Lys	Asn	Thr	Pro	Val	Pro	Ala	
					645					650					655		
35	Asp	Pro	Pro	Leu	Thr	Phe	Asn	Gln	Ala	Lys	Leu	Asn	Ser	Phe	Ile	Thr	
				660					665					670			
	Gln	Tyr	Ser	Thr	Gly	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	Leu	Gln	
40			675					680					685				
	Lys	Glu	Asn	Ser	Lys	Arg	Trp	Asn	Pro	Glu	Ile	Gln	Tyr	Thr	Ser	Asn	
		690					695					700					
45	Tyr	Tyr	Lys	Ser	Thr	Asn	Val	Asp	Phe	Ala	Val	Asn	Thr	Glu	Gly	Val	
	705					710					715					720	
	Tyr	Ser	Glu	Pro	Arg	Pro	Ile	Gly	Thr	Arg	Tyr	Leu	Thr	Arg	Asn	Leu	
50					725					730					735		
	<210>	100															
	<211>	736															
	<212>	PRT															
55	<213>	capsid protein of AAV serotype, clone AAV9															

# EP 1 310 571 A2

400> 100  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

5  
Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

10  
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

15  
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

20  
Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

25  
Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
85 90 95

30  
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
100 105 110

35  
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
115 120 125

40  
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
130 135 140

45  
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly  
145 150 155 160

50  
Lys Ser Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr  
165 170 175

55  
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro  
180 185 190

60  
Glu Ala Pro Ser Gly Leu Gly Pro Asn Thr Met Ala Ser Gly Gly Gly  
195 200 205

65  
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser  
210 215 220

70  
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile  
225 230 235 240

75  
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu  
245 250 255

# EP 1 310 571 A2

Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn  
 260 265 270  
 5 Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg  
 275 280 285  
 10 Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn  
 290 295 300  
 Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile  
 305 310 315 320  
 15 Gln Val Lys Glu Val Thr Thr Asn Glu Gly Thr Lys Thr Ile Ala Asn  
 325 330 335  
 20 Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu  
 340 345 350  
 Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro  
 355 360 365  
 25 Ala Asp Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn  
 370 375 380  
 30 Gly Ser Gln Ala Leu Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe  
 385 390 395 400  
 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr  
 405 410 415  
 35 Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu  
 420 425 430  
 40 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Val  
 435 440 445  
 Arg Thr Gln Thr Thr Gly Thr Gly Gly Thr Gln Thr Leu Ala Phe Ser  
 450 455 460  
 45 Gln Ala Gly Pro Ser Ser Met Ala Asn Gln Ala Arg Asn Trp Val Pro  
 465 470 475 480  
 50 Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Asn Gln Asn  
 485 490 495  
 Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Ala Lys Phe Lys Leu Asn  
 500 505 510  
 55

# EP 1 310 571 A2

	Gly	Arg	Asp	Ser	Leu	Met	Asn	Pro	Gly	Val	Ala	Met	Ala	Ser	His	Lys
			515					520					525			
5	Asp	Asp	Glu	Asp	Arg	Phe	Phe	Pro	Ser	Ser	Gly	Val	Leu	Ile	Phe	Gly
		530					535					540				
	Lys	Gln	Gly	Ala	Gly	Asn	Asp	Gly	Val	Asp	Tyr	Ser	Gln	Val	Leu	Ile
10	545					550					555					560
	Thr	Asp	Glu	Glu	Glu	Ile	Lys	Ala	Thr	Asn	Pro	Val	Ala	Thr	Glu	Glu
					565					570					575	
15	Tyr	Gly	Ala	Val	Ala	Ile	Asn	Asn	Gln	Ala	Ala	Asn	Thr	Gln	Ala	Gln
				580					585					590		
	Thr	Gly	Leu	Val	His	Asn	Gln	Gly	Val	Ile	Pro	Gly	Met	Val	Trp	Gln
20			595					600					605			
	Asn	Arg	Asp	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	Pro	His
		610					615					620				
25	Thr	Asp	Gly	Asn	Phe	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	Gly	Leu
	625					630					635					640
	Lys	His	Pro	Pro	Pro	Gln	Ile	Leu	Ile	Lys	Asn	Thr	Pro	Val	Pro	Ala
30					645					650					655	
	Asp	Pro	Pro	Leu	Thr	Phe	Asn	Gln	Ala	Lys	Leu	Asn	Ser	Phe	Ile	Thr
			660					665						670		
35	Gln	Tyr	Ser	Thr	Gly	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	Leu	Gln
		675						680					685			
	Lys	Glu	Asn	Ser	Lys	Arg	Trp	Asn	Pro	Glu	Ile	Gln	Tyr	Thr	Ser	Asn
40		690					695					700				
	Tyr	Tyr	Lys	Ser	Thr	Asn	Val	Asp	Phe	Ala	Val	Asn	Thr	Glu	Gly	Val
	705					710					715					720
45	Tyr	Ser	Glu	Pro	Arg	Pro	Ile	Gly	Thr	Arg	Tyr	Leu	Thr	Arg	Asn	Leu
					725					730					735	
50	<210>	101														
	<211>	728														
	<212>	PRT														
	<213>	capsid protein of AAV serotype, clone 24.1														
55	<400>	101														
	Met	Ala	Ala	Asp	Gly	Tyr	Leu	Pro	Asp	Trp	Leu	Glu	Asp	Asn	Leu	Ser
	1				5					10					15	

# EP 1 310 571 A2

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 5  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 10  
 Gly Tyr Lys Tyr Leu Arg Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 15  
 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 20  
 Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 85 90 95  
 25  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 30  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 35  
 Leu Gly Leu Val Glu Glu Val Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 40  
 Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln  
 145 150 155 160  
 45  
 Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu  
 165 170 175  
 50  
 Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser  
 180 185 190  
 55  
 Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala  
 195 200 205  
 60  
 Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp  
 210 215 220  
 65  
 His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr  
 225 230 235 240  
 70  
 Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile  
 245 250 255  
 75  
 Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Ser Tyr Ser  
 260 265 270

# EP 1 310 571 A2

	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg	Phe	His	Cys	His	Phe	Ser
			275					280					285			
5																
	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn	Asn	Trp	Gly	Phe	Arg	Pro
		290					295					300				
10	Arg	Lys	Leu	Arg	Phe	Lys	Leu	Phe	Asn	Ile	Gln	Val	Lys	Glu	Val	Thr
	305					310					315					320
	Thr	Asn	Asp	Gly	Val	Thr	Thr	Ile	Ala	Asn	Asn	Leu	Thr	Ser	Thr	Ile
					325					330					335	
15																
	Gln	Val	Phe	Ser	Asp	Ser	Glu	Tyr	Gln	Leu	Pro	Tyr	Val	Leu	Gly	Ser
				340					345					350		
20	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	Pro	Ala	Asp	Val	Phe	Met	Ile
			355					360					365			
	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asn	Gly	Ser	Gln	Ser	Val	Gly
	370						375					380				
25																
	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe	Pro	Ser	Gln	Met	Leu	Arg
	385					390					395					400
30	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	Thr	Phe	Glu	Glu	Val	Pro	Phe
					405					410					415	
	His	Ser	Ser	Tyr	Val	His	Ser	Gln	Ser	Leu	Asp	Arg	Leu	Met	Asn	Pro
				420					425					430		
35																
	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ala	Arg	Thr	Gln	Ser	Thr	Thr
			435					440					445			
40	Gly	Ser	Thr	Arg	Glu	Leu	Gln	Phe	His	Gln	Ala	Gly	Pro	Asn	Thr	Met
	450						455					460				
	Ala	Glu	Gln	Ser	Lys	Asn	Trp	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln
	465					470					475					480
45																
	Arg	Leu	Ser	Lys	Asn	Ile	Asp	Ser	Asn	Asn	Asn	Ser	Asn	Phe	Ala	Trp
					485					490					495	
50	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly	Arg	Asn	Ser	Leu	Thr	Asn
				500					505					510		
	Pro	Gly	Val	Ala	Met	Ala	Thr	Asn	Lys	Asp	Asp	Glu	Asp	Gln	Phe	Phe
			515					520					525			
55																

# EP 1 310 571 A2

5 Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn Lys  
 530 535 540  
 Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys Thr  
 545 550 555 560  
 10 Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn Leu  
 565 570 575  
 Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly  
 580 585 590  
 15 Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Cys Leu Gln Gly  
 595 600 605  
 20 Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser  
 610 615 620  
 Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu  
 625 630 635 640  
 25 Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr Pro  
 645 650 655  
 30 Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser  
 660 665 670  
 Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn  
 675 680 685  
 35 Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu  
 690 695 700  
 40 Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly  
 705 710 715 720  
 Thr Arg Tyr Leu Thr Arg Asn Leu  
 725  
 45  
 <210> 102  
 <211> 728  
 <212> PRT  
 50 <213> capsid protein of AAV serotype, clone 42.2REAL  
 <400> 102  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 55

# EP 1 310 571 A2

	Glu	Gly	Ile	Arg	Glu	Trp	Trp	Asp	Leu	Lys	Pro	Gly	Ala	Pro	Lys	Pro	
				20					25					30			
5	Lys	Ala	Asn	Gln	Gln	Lys	Gln	Asp	Asp	Gly	Arg	Gly	Leu	Val	Leu	Pro	
			35					40					45				
10	Gly	Tyr	Lys	Tyr	Leu	Gly	Pro	Phe	Asn	Gly	Leu	Asp	Lys	Gly	Glu	Pro	
		50					55					60					
15	Val	Asn	Glu	Ala	Asp	Ala	Ala	Ala	Leu	Glu	His	Asp	Lys	Ala	Tyr	Asp	
	65					70					75					80	
20	Lys	Gln	Leu	Glu	Gln	Gly	Asp	Asn	Pro	Tyr	Leu	Lys	Tyr	Asn	His	Ala	
					85					90					95		
25	Asp	Ala	Glu	Phe	Gln	Glu	Arg	Leu	Gln	Glu	Asp	Thr	Ser	Phe	Gly	Gly	
				100					105						110		
30	Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Val	Leu	Glu	Pro	
			115					120					125				
35	Leu	Gly	Leu	Val	Glu	Glu	Gly	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg	
		130					135					140					
40	Pro	Ile	Glu	Ser	Pro	Asp	Ser	Ser	Thr	Gly	Ile	Gly	Lys	Lys	Gly	Gln	
	145					150					155					160	
45	Gln	Pro	Ala	Lys	Lys	Lys	Leu	Asn	Phe	Gly	Gln	Thr	Gly	Asp	Ser	Glu	
					165					170					175		
50	Ser	Val	Pro	Asp	Pro	Gln	Pro	Leu	Gly	Glu	Pro	Pro	Ala	Ala	Pro	Ser	
				180					185					190			
55	Gly	Leu	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly	Gly	Ala	Pro	Met	Ala	
			195					200					205				
60	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Asn	Ala	Ser	Gly	Asn	Trp	
		210					215					220					
65	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val	Ile	Thr	Thr	Ser	Thr	
	225					230					235					240	
70	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	Leu	Tyr	Lys	Gln	Ile	
					245					250					255		
75	Ser	Ser	Gln	Ser	Gly	Ala	Thr	Asn	Asp	Asn	His	Phe	Phe	Gly	Tyr	Ser	
				260					265					270			

EP 1 310 571 A2

	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg	Phe	His	Cys	His	Phe	Ser
			275					280					285			
5	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn	Asn	Trp	Gly	Phe	Arg	Pro
		290					295					300				
10	Arg	Lys	Leu	Arg	Phe	Lys	Leu	Phe	Asn	Ile	Gln	Val	Lys	Glu	Val	Thr
	305					310					315					320
	Thr	Asn	Asp	Gly	Val	Thr	Thr	Ile	Ala	Asn	Asn	Leu	Thr	Ser	Thr	Ile
					325					330					335	
15	Gln	Val	Phe	Ser	Asp	Ser	Glu	Tyr	Gln	Leu	Pro	Tyr	Val	Leu	Gly	Ser
				340					345					350		
20	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	Pro	Ala	Asp	Val	Phe	Met	Ile
			355					360					365			
	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asn	Gly	Ser	Gln	Ser	Val	Gly
		370					375					380				
25	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe	Pro	Ser	Gln	Met	Leu	Arg
	385					390					395					400
30	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	Thr	Phe	Glu	Glu	Val	Pro	Phe
					405					410					415	
	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	Leu	Asp	Arg	Leu	Met	Asn	Pro
				420					425					430		
35	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ala	Arg	Thr	Gln	Ser	Thr	Thr
			435					440					445			
40	Gly	Ser	Thr	Arg	Glu	Leu	Gln	Phe	His	Gln	Ala	Gly	Pro	Asn	Thr	Met
		450					455					460				
45	Ala	Glu	Gln	Ser	Lys	Asn	Trp	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln
	465					470					475					480
	Arg	Leu	Ser	Lys	Asn	Ile	Asp	Ser	Asn	Asn	Asn	Ser	Asn	Phe	Ala	Trp
				485						490					495	
50	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly	Arg	Asn	Ser	Leu	Thr	Asn
				500					505					510		
55	Pro	Gly	Val	Ala	Met	Ala	Thr	Asn	Lys	Asp	Asp	Glu	Asp	Gln	Phe	Phe
			515					520					525			

# EP 1 310 571 A2

	Pro	Ile	Asn	Gly	Val	Leu	Val	Phe	Gly	Glu	Thr	Gly	Ala	Ala	Asn	Lys	
	530						535					540					
5	Thr	Thr	Leu	Glu	Asn	Val	Leu	Met	Thr	Ser	Glu	Glu	Glu	Ile	Lys	Thr	
	545					550					555					560	
	Thr	Asn	Pro	Val	Ala	Thr	Glu	Glu	Tyr	Gly	Val	Val	Ser	Ser	Asn	Leu	
10					565					570					575		
	Gln	Ser	Ser	Thr	Ala	Gly	Pro	Gln	Thr	Gln	Thr	Val	Asn	Ser	Gln	Gly	
				580					585					590			
15	Ala	Leu	Pro	Gly	Met	Val	Trp	Gln	Asn	Arg	Asp	Val	Tyr	Leu	Gln	Gly	
			595					600					605				
	Pro	Ile	Trp	Ala	Lys	Ile	Pro	His	Thr	Asp	Gly	Asn	Phe	His	Pro	Ser	
20		610					615					620					
	Pro	Leu	Met	Gly	Gly	Phe	Gly	Leu	Lys	His	Pro	Pro	Pro	Gln	Ile	Leu	
	625					630					635					640	
25	Ile	Lys	Asn	Thr	Pro	Val	Pro	Ala	Asn	Pro	Pro	Glu	Val	Phe	Thr	Pro	
					645					650					655		
	Ala	Lys	Phe	Ala	Ser	Phe	Ile	Thr	Gln	Tyr	Ser	Thr	Gly	Gln	Val	Ser	
30				660					665					670			
	Val	Glu	Ile	Glu	Trp	Glu	Leu	Gln	Lys	Glu	Asn	Ser	Lys	Arg	Trp	Asn	
			675					680					685				
35	Pro	Glu	Ile	Gln	Tyr	Thr	Ser	Asn	Tyr	Ala	Lys	Ser	Asn	Asn	Val	Glu	
		690					695					700					
	Phe	Ala	Val	Asn	Asn	Glu	Gly	Val	Tyr	Thr	Glu	Pro	Arg	Pro	Ile	Gly	
40	705					710					715					720	
	Thr	Arg	Tyr	Leu	Thr	Arg	Asn	Leu									
					725												
45	<210>	103															
	<211>	728															
	<212>	PRT															
	<213>	capsid protein of AAV serotype, clone 7.2VP1															
50	<400>	103															
	Met	Ala	Ala	Asp	Gly	Tyr	Leu	Pro	Asp	Trp	Leu	Glu	Gly	Asn	Leu	Ser	
	1				5					10					15		
55	Glu	Gly	Ile	Arg	Glu	Trp	Trp	Asp	Leu	Lys	Pro	Gly	Ala	Pro	Lys	Pro	
				20					25					30			

# EP 1 310 571 A2

5 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 Gly Tyr Arg Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 10 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 85 90 95  
 15 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 20 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 25 Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Asn Gly Gln  
 145 150 155 160  
 30 Pro Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu  
 165 170 175  
 Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser  
 180 185 190  
 Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala  
 195 200 205  
 40 Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp  
 210 215 220  
 His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr  
 225 230 235 240  
 Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile  
 245 250 255  
 50 Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser  
 260 265 270  
 Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser  
 275 280 285

# EP 1 310 571 A2

	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn	Asn	Trp	Gly	Phe	Arg	Pro	
	290						295					300					
5																	
	Arg	Lys	Leu	Arg	Phe	Lys	Leu	Phe	Asn	Ile	Gln	Val	Lys	Glu	Val	Thr	
	305					310					315					320	
10	Thr	Asn	Asp	Gly	Val	Thr	Thr	Ile	Ala	Asn	Asn	Leu	Thr	Ser	Thr	Ile	
					325					330					335		
	Gln	Val	Phe	Ser	Asp	Ser	Glu	Tyr	Gln	Leu	Pro	Tyr	Val	Leu	Gly	Ser	
				340					345					350			
15																	
	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	Pro	Ala	Asp	Val	Phe	Met	Ile	
			355					360					365				
20	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asn	Gly	Ser	Gln	Ser	Val	Gly	
	370						375					380					
	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe	Pro	Ser	Gln	Met	Leu	Arg	
	385					390					395					400	
25																	
	Thr	Gly	Asp	Asn	Phe	Glu	Phe	Ser	Tyr	Thr	Phe	Glu	Glu	Val	Pro	Phe	
					405					410					415		
30	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	Leu	Asp	Arg	Leu	Met	Asn	Pro	
				420					425					430			
	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ala	Arg	Thr	Gln	Ser	Thr	Thr	
			435					440					445				
35																	
	Gly	Ser	Thr	Arg	Glu	Leu	Gln	Phe	His	Gln	Ala	Gly	Pro	Asn	Thr	Met	
	450						455					460					
40	Ala	Glu	Gln	Ser	Lys	Asn	Trp	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	
	465					470					475					480	
	Arg	Leu	Ser	Lys	Asn	Ile	Asp	Ser	Asn	Asn	Asn	Ser	Asn	Phe	Ala	Trp	
					485					490					495		
45																	
	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly	Arg	Asn	Ser	Leu	Thr	Asn	
					500				505					510			
50	Pro	Gly	Val	Ala	Met	Ala	Thr	Asn	Lys	Asp	Asp	Glu	Asp	Gln	Phe	Phe	
			515					520					525				
	Pro	Ile	Asn	Gly	Val	Leu	Val	Phe	Gly	Lys	Thr	Gly	Ala	Ala	Asn	Lys	
	530						535					540					
55																	

# EP 1 310 571 A2

Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys Thr  
 545 550 555 560  
 5  
 Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn Leu  
 565 570 575  
 10  
 Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly  
 580 585 590  
 Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly  
 595 600 605  
 15  
 Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser  
 610 615 620  
 20  
 Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu  
 625 630 635 640  
 Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr Pro  
 645 650 655  
 25  
 Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser  
 660 665 670  
 30  
 Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn  
 675 680 685  
 Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu  
 690 695 700  
 35  
 Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly  
 705 710 715 720  
 40  
 Thr Arg Tyr Leu Thr Arg Asn Leu  
 725  
 45  
 <210> 104  
 <211> 728  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 27.3VP1  
 <400> 104  
 50  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 55

# EP 1 310 571 A2

	Lys	Ala	Asn	Gln	Gln	Lys	Gln	Asp	Asp	Gly	Arg	Gly	Leu	Val	Leu	Pro
			35					40					45			
5	Gly	Tyr	Lys	Tyr	Leu	Gly	Pro	Phe	Asn	Gly	Leu	Asp	Lys	Gly	Glu	Pro
		50					55					60				
	Val	Asn	Glu	Ala	Asp	Ala	Ala	Ala	Leu	Glu	His	Asp	Lys	Ala	Tyr	Asp
10	65					70					75					80
	Lys	Gln	Leu	Glu	Gln	Gly	Asp	Asn	Pro	Tyr	Leu	Lys	Tyr	Asn	His	Ala
					85					90					95	
15	Asp	Ala	Glu	Phe	Gln	Glu	Arg	Leu	Gln	Glu	Asp	Thr	Ser	Phe	Gly	Gly
				100					105					110		
	Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Val	Leu	Glu	Pro
20			115					120					125			
	Leu	Gly	Leu	Val	Glu	Glu	Gly	Ala	Lys	Thr	Ala	Ser	Gly	Lys	Lys	Arg
		130					135					140				
25	Pro	Ile	Glu	Ser	Pro	Asp	Ser	Ser	Thr	Gly	Ile	Gly	Lys	Lys	Gly	Gln
	145					150					155					160
	Gln	Pro	Ala	Lys	Lys	Lys	Leu	Asn	Phe	Gly	Gln	Thr	Gly	Asp	Ser	Glu
30					165					170					175	
	Ser	Val	Pro	Asp	Pro	Gln	Pro	Leu	Gly	Glu	Pro	Pro	Ala	Ala	Pro	Ser
				180					185					190		
35	Gly	Leu	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly	Gly	Ala	Pro	Met	Ala
		195						200					205			
	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Asn	Ala	Ser	Gly	Asn	Trp
40		210					215					220				
	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val	Ile	Thr	Thr	Ser	Thr
	225					230					235					240
45	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	Leu	Tyr	Lys	Gln	Ile
				245					250						255	
	Ser	Ser	Gln	Ser	Gly	Ala	Thr	Asn	Asp	Asn	His	Phe	Phe	Gly	Tyr	Ser
50				260					265					270		
	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg	Phe	His	Cys	His	Phe	Ser
			275					280					285			
55																

# EP 1 310 571 A2

	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn	Asn	Trp	Gly	Phe	Arg	Pro	
	290						295					300					
5	Arg	Lys	Leu	Arg	Phe	Lys	Leu	Phe	Asn	Ile	Gln	Val	Lys	Glu	Val	Thr	
	305					310					315					320	
	Thr	Asn	Asp	Gly	Val	Thr	Thr	Ile	Ala	Asn	Asn	Leu	Thr	Ser	Thr	Ile	
10					325					330					335		
	Gln	Val	Phe	Ser	Asp	Ser	Glu	Tyr	Gln	Leu	Pro	Tyr	Val	Leu	Gly	Ser	
				340					345					350			
15	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	Pro	Ala	Asp	Val	Phe	Met	Ile	
			355					360					365				
	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asn	Gly	Ser	Gln	Ser	Val	Gly	
20		370					375					380					
	Arg	Ser	Ser	Phe	Cys	Cys	Leu	Glu	Tyr	Phe	Pro	Ser	Gln	Met	Leu	Arg	
	385					390					395					400	
25	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	Thr	Phe	Glu	Glu	Val	Pro	Phe	
					405					410					415		
	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	Leu	Asp	Arg	Leu	Met	Asn	Pro	
30				420					425					430			
	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ala	Arg	Thr	Gln	Ser	Thr	Thr	
			435					440					445				
35	Gly	Ser	Thr	Arg	Glu	Leu	Gln	Phe	His	Gln	Ala	Gly	Pro	Asn	Thr	Val	
	450						455					460					
	Ala	Glu	Gln	Ser	Lys	Asn	Trp	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	
40	465					470					475					480	
	Arg	Leu	Ser	Lys	Asn	Ile	Asp	Ser	Asn	Asn	Asn	Ser	Asn	Phe	Ala	Trp	
				485						490				495			
45	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly	Arg	Asn	Ser	Leu	Thr	Asn	
				500					505					510			
	Pro	Gly	Val	Ala	Met	Ala	Thr	Asn	Lys	Asp	Asp	Glu	Asp	Gln	Phe	Leu	
50			515					520					525				
	Pro	Ile	Asn	Gly	Val	Leu	Val	Phe	Gly	Lys	Thr	Gly	Ala	Ala	Asn	Lys	
	530						535					540					
55																	

# EP 1 310 571 A2

Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys Thr  
 545 550 555 560  
 5  
 Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn Leu  
 565 570 575  
 10  
 Gln Ser Ser Thr Ala Gly Pro Arg Thr Gln Thr Val Asn Ser Gln Gly  
 580 585 590  
 Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly  
 595 600 605  
 15  
 Pro Ile Trp Ala Glu Ile Pro His Thr Asp Gly Asn Phe His Pro Ser  
 610 615 620  
 20  
 Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu  
 625 630 635 640  
 Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr Pro  
 645 650 655  
 25  
 Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser  
 660 665 670  
 30  
 Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn  
 675 680 685  
 Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu  
 690 695 700  
 35  
 Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly  
 705 710 715 720  
 40  
 Thr Arg Tyr Leu Thr Arg Asn Leu  
 725  
 45  
 <210> 105  
 <211> 728  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 16.3VP1  
 <400> 105  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 50  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 55  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45

# EP 1 310 571 A2

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 5  
 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 10  
 Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 85 90 95  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 15  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 20  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln  
 145 150 155 160  
 25  
 Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu  
 165 170 175  
 30  
 Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser  
 180 185 190  
 Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala  
 195 200 205  
 35  
 Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp  
 210 215 220  
 40  
 His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr  
 225 230 235 240  
 Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile  
 245 250 255  
 45  
 Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser  
 260 265 270  
 50  
 Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser  
 275 280 285  
 Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro  
 290 295 300  
 55

# EP 1 310 571 A2

Arg Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr  
 305 310 315 320  
 5  
 Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile  
 325 330 335  
 10  
 Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser  
 340 345 350  
 Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile  
 355 360 365  
 15  
 Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Met Gly  
 370 375 380  
 20  
 Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg  
 385 390 395 400  
 Thr Gly Asn Asn Phe Glu Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe  
 405 410 415  
 25  
 His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro  
 420 425 430  
 30  
 Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr Thr  
 435 440 445  
 Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr Met  
 450 455 460  
 35  
 Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln  
 465 470 475 480  
 40  
 Arg Leu Ser Lys Asn Ile Asp Ser Asn Asn Asn Ser Asn Phe Ala Trp  
 485 490 495  
 Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr Asn  
 500 505 510  
 45  
 Pro Gly Val Ala Met Ala Thr Asn Lys Asp Asp Glu Gly Gln Phe Phe  
 515 520 525  
 50  
 Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn Lys  
 530 535 540  
 Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys Thr  
 545 550 555 560  
 55

# EP 1 310 571 A2

Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn Leu  
 565 570 575  
 5  
 Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly  
 580 585 590  
 10  
 Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly  
 595 600 605  
 Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser  
 610 615 620  
 15  
 Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu  
 625 630 635 640  
 20  
 Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Gly Val Phe Thr Pro  
 645 650 655  
 Ala Leu Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser  
 660 665 670  
 25  
 Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn  
 675 680 685  
 30  
 Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu  
 690 695 700  
 Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly  
 705 710 715 720  
 35  
 Thr Arg Tyr Leu Thr Arg Asn Leu  
 725  
 40  
 <210> 106  
 <211> 728  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 42.10  
 <400> 106  
 45  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 50  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 55

# EP 1 310 571 A2

	Gly	Tyr	Lys	Tyr	Leu	Gly	Pro	Phe	Asn	Gly	Leu	Asp	Lys	Gly	Glu	Pro	
	50						55					60					
5	Val	Asn	Glu	Ala	Asp	Ala	Ala	Ala	Leu	Glu	His	Asp	Lys	Ala	Tyr	Asp	
	65					70					75					80	
	Lys	Gln	Leu	Glu	Gln	Gly	Asp	Asn	Pro	Tyr	Leu	Lys	Tyr	Asn	His	Ala	
10					85					90					95		
	Asp	Ala	Glu	Phe	Gln	Glu	Arg	Leu	Gln	Glu	Asp	Thr	Ser	Phe	Gly	Gly	
				100					105						110		
15	Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Val	Leu	Glu	Pro	
			115					120					125				
	Leu	Gly	Leu	Val	Glu	Glu	Gly	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg	
20		130					135					140					
	Pro	Ile	Glu	Ser	Pro	Asp	Ser	Ser	Thr	Gly	Ile	Gly	Arg	Lys	Gly	Gln	
	145					150					155					160	
25	Gln	Pro	Ala	Lys	Lys	Lys	Leu	Asn	Phe	Gly	Gln	Thr	Gly	Asp	Ser	Glu	
					165					170					175		
	Ser	Val	Pro	Asp	Pro	Gln	Pro	Ile	Gly	Glu	Pro	Pro	Ala	Gly	Pro	Ser	
30				180					185					190			
	Gly	Leu	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly	Gly	Ala	Pro	Met	Ala	
			195					200					205				
35	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Asn	Ala	Ser	Gly	Asn	Trp	
		210					215					220					
	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val	Ile	Thr	Thr	Ser	Thr	
40		225				230					235					240	
	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	Leu	Tyr	Lys	Gln	Ile	
					245					250					255		
45	Ser	Ser	Gln	Ser	Gly	Ala	Thr	Asn	Asp	Asn	His	Phe	Phe	Gly	Tyr	Ser	
				260					265					270			
	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg	Phe	His	Cys	His	Phe	Ser	
50			275					280					285				
	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn	Asn	Trp	Gly	Phe	Arg	Pro	
		290					295					300					
55																	

# EP 1 310 571 A2

	Arg	Lys	Leu	Arg	Phe	Lys	Leu	Phe	Asn	Ile	Gln	Val	Lys	Glu	Val	Thr	305	310	315	320
5	Thr	Asn	Asp	Gly	Val	Thr	Thr	Ile	Ala	Asn	Asn	Leu	Thr	Ser	Thr	Ile	325	330	335	
10	Gln	Val	Phe	Ser	Asp	Ser	Glu	Tyr	Gln	Leu	Pro	Tyr	Val	Leu	Gly	Ser	340	345	350	
15	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	Pro	Ala	Asp	Val	Phe	Met	Ile	355	360	365	
20	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asn	Gly	Ser	Gln	Ser	Val	Gly	370	375	380	
25	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe	Pro	Ser	Gln	Met	Leu	Arg	385	390	395	400
30	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	Thr	Phe	Glu	Glu	Val	Pro	Phe	405	410	415	
35	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	Leu	Asp	Arg	Leu	Met	Asn	Pro	420	425	430	
40	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ala	Arg	Thr	Gln	Ser	Thr	Thr	435	440	445	
45	Gly	Ser	Thr	Arg	Glu	Leu	Gln	Phe	His	Gln	Ala	Gly	Pro	Asn	Thr	Met	450	455	460	
50	Ala	Glu	Gln	Ser	Lys	Asn	Trp	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	465	470	475	480
55	Arg	Leu	Ser	Lys	Asn	Ile	Asp	Ser	Asn	Asn	Asn	Ser	Asn	Phe	Ala	Trp	485	490	495	
	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly	Arg	Asn	Ser	Leu	Thr	Asn	500	505	510	
	Pro	Gly	Val	Ala	Met	Ala	Thr	Asn	Lys	Asp	Asp	Glu	Asp	Gln	Phe	Phe	515	520	525	
	Pro	Ile	Asn	Gly	Val	Leu	Val	Phe	Gly	Lys	Thr	Gly	Ala	Ala	Asn	Lys	530	535	540	
	Thr	Thr	Leu	Glu	Asn	Val	Leu	Met	Thr	Ser	Glu	Glu	Glu	Ile	Lys	Thr	545	550	555	560

# EP 1 310 571 A2

	Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn Leu	565	570	575
5	Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly	580	585	590
10	Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly	595	600	605
	Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser	610	615	620
15	Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu	625	630	635
	Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr Pro	645	650	655
20	Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser	660	665	670
25	Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn	675	680	685
30	Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu	690	695	700
	Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly	705	710	715
35	Thr Arg Tyr Leu Thr Arg Asn Leu	725		
40	<210> 107 <211> 728 <212> PRT <213> capsid protein of AAV serotype, clone 42.3B			
45	<400> 107 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser	1	5	10
	Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro	20	25	30
50	Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro	35	40	45
55	Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro	50	55	60

# EP 1 310 571 A2

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 5  
 Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 85 90 95  
 10  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 15  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 20  
 Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln  
 145 150 155 160  
 Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu  
 165 170 175  
 25  
 Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Gly Pro Ser  
 180 185 190  
 30  
 Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala  
 195 200 205  
 Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp  
 210 215 220  
 35  
 His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr  
 225 230 235 240  
 40  
 Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile  
 245 250 255  
 Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser  
 260 265 270  
 45  
 Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser  
 275 280 285  
 50  
 Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro  
 290 295 300  
 Arg Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr  
 305 310 315 320  
 55

# EP 1 310 571 A2

Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile  
 325 330 335  
 5  
 Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser  
 340 345 350  
 10  
 Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile  
 355 360 365  
 Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val Gly  
 370 375 380  
 15  
 Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg  
 385 390 395 400  
 20  
 Thr Gly Asn Asn Phe Glu Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe  
 405 410 415  
 His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro  
 420 425 430  
 25  
 Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr Thr  
 435 440 445  
 30  
 Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr Met  
 450 455 460  
 Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln  
 465 470 475 480  
 35  
 Arg Leu Ser Lys Asn Ile Asp Ser Asn Asn Thr Ser Asn Phe Ala Trp  
 485 490 495  
 40  
 Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr Asn  
 500 505 510  
 Pro Gly Val Ala Met Ala Thr Asn Lys Asp Asp Glu Asp Gln Phe Phe  
 515 520 525  
 45  
 Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn Lys  
 530 535 540  
 50  
 Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys Thr  
 545 550 555 560  
 Thr Asn Pro Val Ala Thr Glu Gln Tyr Gly Val Val Ser Ser Asn Leu  
 565 570 575  
 55

# EP 1 310 571 A2

Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly  
 580 585 590  
 5  
 Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly  
 595 600 605  
 10  
 Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser  
 610 615 620  
 Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu  
 625 630 635 640  
 15  
 Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr Pro  
 645 650 655  
 20  
 Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser  
 660 665 670  
 Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn  
 675 680 685  
 25  
 Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu  
 690 695 700  
 30  
 Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly  
 705 710 715 720  
 Thr Arg Tyr Leu Thr Arg Asn Leu  
 725  
 35  
 <210> 108  
 <211> 728  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 42.11  
 40  
 <400> 108  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 45  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 50  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 55

# EP 1 310 571 A2

	Val	Asn	Ala	Ala	Asp	Ala	Ala	Ala	Leu	Glu	His	Asp	Lys	Ala	Tyr	Asp	65	70	75	80
5	Gln	Gln	Leu	Lys	Ala	Gly	Asp	Asn	Pro	Tyr	Leu	Arg	Tyr	Asn	His	Ala	85	90	95	
10	Asp	Ala	Glu	Phe	Gln	Glu	Arg	Leu	Gln	Glu	Asp	Thr	Ser	Phe	Gly	Gly	100	105	110	
	Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Val	Leu	Glu	Pro	115	120	125	
15	Leu	Gly	Leu	Val	Glu	Glu	Gly	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg	130	135	140	
20	Pro	Ile	Glu	Ser	Pro	Asp	Ser	Ser	Thr	Gly	Ile	Gly	Lys	Lys	Gly	Gln	145	150	155	160
	Gln	Pro	Ala	Lys	Lys	Lys	Leu	Asn	Phe	Gly	Gln	Thr	Gly	Asp	Ser	Glu	165	170	175	
25	Ser	Val	Pro	Asp	Pro	Gln	Pro	Ile	Gly	Glu	Pro	Pro	Ala	Gly	Pro	Ser	180	185	190	
30	Gly	Leu	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly	Gly	Ala	Pro	Met	Ala	195	200	205	
	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Asn	Ala	Ser	Gly	Asn	Trp	210	215	220	
35	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val	Ile	Thr	Thr	Ser	Thr	225	230	235	240
40	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	Leu	Tyr	Lys	Gln	Ile	245	250	255	
	Ser	Ser	Gln	Ser	Gly	Ala	Thr	Asn	Asp	Asn	His	Phe	Phe	Gly	Tyr	Ser	260	265	270	
45	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg	Phe	His	Cys	His	Phe	Ser	275	280	285	
50	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn	Asn	Trp	Gly	Phe	Arg	Pro	290	295	300	
	Arg	Lys	Leu	Arg	Phe	Lys	Leu	Phe	Asn	Ile	Gln	Val	Lys	Glu	Val	Thr	305	310	315	320
55																				

# EP 1 310 571 A2

	Thr	Asn	Asp	Gly	Val	Thr	Thr	Ile	Ala	Asn	Asn	Leu	Thr	Ser	Thr	Ile	
					325					330					335		
5	Gln	Val	Phe	Ser	Asp	Ser	Glu	Tyr	Gln	Leu	Pro	Tyr	Val	Leu	Gly	Ser	
				340					345					350			
	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	Pro	Ala	Asp	Val	Phe	Met	Ile	
10			355					360					365				
	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asn	Gly	Ser	Gln	Ser	Val	Gly	
		370					375					380					
15	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe	Pro	Ser	Gln	Met	Leu	Arg	
	385					390					395					400	
	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	Thr	Phe	Glu	Glu	Val	Pro	Phe	
20					405					410					415		
	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	Leu	Asp	Arg	Leu	Met	Asn	Pro	
				420					425					430			
25	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ala	Arg	Thr	Gln	Ser	Thr	Thr	
			435					440					445				
	Gly	Ser	Thr	Arg	Glu	Leu	Gln	Phe	His	Gln	Ala	Gly	Pro	Asn	Thr	Met	
30		450					455					460					
	Ala	Glu	Gln	Ser	Lys	Asn	Trp	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Arg	Gln	
	465					470					475					480	
35	Arg	Leu	Ser	Lys	Asp	Ile	Asp	Ser	Asn	Asn	Asn	Ser	Asn	Phe	Ala	Trp	
					485					490					495		
	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly	Arg	Asn	Ser	Leu	Thr	Asn	
40				500					505					510			
	Pro	Gly	Val	Ala	Met	Ala	Thr	Asn	Lys	Asp	Asp	Glu	Asp	Gln	Phe	Phe	
			515					520					525				
45	Pro	Ile	Asn	Gly	Val	Leu	Val	Phe	Gly	Lys	Thr	Gly	Ala	Ala	Asn	Lys	
		530					535					540					
	Thr	Thr	Leu	Glu	Asn	Val	Leu	Met	Thr	Ser	Glu	Glu	Glu	Ile	Lys	Thr	
50		545				550					555					560	
	Thr	Asn	Pro	Val	Ala	Thr	Glu	Glu	Tyr	Gly	Val	Val	Ser	Ser	Asn	Leu	
					565					570					575		
55																	

# EP 1 310 571 A2

Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly  
580 585 590

5 Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly  
595 600 605

10 Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser  
610 615 620

Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu  
625 630 635 640

15 Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr Pro  
645 650 655

20 Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser  
660 665 670

Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn  
675 680 685

25 Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu  
690 695 700

30 Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly  
705 710 715 720

Thr Arg Tyr Leu Thr Arg Asn Leu  
725

35 <210> 109  
<211> 729  
<212> PRT  
<213> capsid protein of AAV serotype, clone FlVP1

40 <400> 109

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

45 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

50 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

55

# EP 1 310 571 A2

	Gln	Gln	Leu	Lys	Ala	Gly	Asp	Asn	Pro	Tyr	Leu	Arg	Tyr	Asn	His	Ala	
					85					90					95		
5	Asp	Ala	Glu	Phe	Gln	Glu	Arg	Leu	Gln	Glu	Asp	Thr	Ser	Phe	Gly	Gly	
				100				105						110			
	Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Val	Leu	Glu	Pro	
10			115					120					125				
	Leu	Gly	Leu	Val	Glu	Glu	Gly	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg	
		130					135					140					
15	Pro	Ile	Asp	Ser	Pro	Asp	Ser	Ser	Thr	Gly	Ile	Gly	Lys	Lys	Gly	Gln	
	145					150					155					160	
	Gln	Pro	Ala	Lys	Lys	Lys	Leu	Asn	Phe	Gly	Gln	Thr	Gly	Asp	Ser	Glu	
20					165					170					175		
	Ser	Val	Pro	Asp	Pro	Gln	Pro	Leu	Gly	Glu	Pro	Pro	Ala	Ala	Pro	Ser	
				180					185					190			
25	Ser	Val	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly	Gly	Ala	Pro	Met	Ala	
			195					200					205				
	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Asn	Ala	Ser	Gly	Asn	Trp	
30		210					215					220					
	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val	Ile	Thr	Thr	Ser	Thr	
	225					230					235					240	
35	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	Leu	Tyr	Lys	Gln	Ile	
					245					250					255		
	Ser	Ser	Ser	Ser	Ser	Gly	Ala	Thr	Asn	Asp	Asn	His	Tyr	Phe	Gly	Tyr	
40					260				265					270			
	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg	Phe	His	Cys	His	Phe	
			275					280					285				
45	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn	Asn	Trp	Gly	Phe	Arg	
		290					295					300					
	Pro	Lys	Lys	Leu	Arg	Phe	Lys	Leu	Phe	Asn	Ile	Gln	Val	Lys	Glu	Val	
50		305				310					315					320	
	Thr	Thr	Asn	Asp	Gly	Val	Thr	Thr	Ile	Ala	Asn	Asn	Leu	Thr	Ser	Thr	
					325					330					335		
55																	

# EP 1 310 571 A2

	Val	Gln	Val	Phe	Ser	Asp	Ser	Glu	Tyr	Gln	Leu	Pro	Tyr	Val	Leu	Gly
				340					345					350		
5	Ser	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	Pro	Ala	Asp	Val	Phe	Met
			355					360					365			
	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asn	Gly	Ser	Gln	Ser	Val
10		370					375					380				
	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe	Pro	Ser	Gln	Met	Leu
	385					390					395					400
15	Arg	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	Ser	Phe	Glu	Asp	Val	Pro
				405						410					415	
	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	Leu	Asp	Arg	Leu	Met	Asn
20				420					425					430		
	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ala	Arg	Thr	Gln	Ser	Thr
			435					440						445		
25	Thr	Gly	Ser	Thr	Arg	Glu	Leu	Gln	Phe	His	Gln	Ala	Gly	Pro	Asn	Thr
	450						455					460				
	Met	Ala	Glu	Gln	Ser	Lys	Asn	Trp	Leu	Pro	Gly	Pro	Cys	Tyr	Arg	Gln
30	465					470					475					480
	Gln	Gly	Leu	Ser	Lys	Asn	Leu	Asp	Phe	Asn	Asn	Asn	Ser	Asn	Phe	Ala
					485					490					495	
35	Trp	Thr	Ala	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly	Arg	Asn	Ser	Leu	Thr
				500					505					510		
	Asn	Pro	Gly	Ile	Pro	Met	Ala	Thr	Asn	Lys	Asp	Asp	Glu	Asp	Gln	Phe
40			515					520					525			
	Phe	Pro	Ile	Asn	Gly	Val	Leu	Val	Phe	Gly	Lys	Thr	Gly	Ala	Ala	Asn
	530						535					540				
45	Lys	Thr	Thr	Leu	Glu	Asn	Val	Leu	Met	Thr	Ser	Glu	Glu	Glu	Ile	Lys
	545					550					555					560
	Thr	Thr	Asn	Pro	Val	Ala	Thr	Glu	Glu	Tyr	Gly	Val	Val	Ser	Ser	Asn
50					565					570					575	
	Leu	Gln	Pro	Ser	Thr	Ala	Gly	Pro	Gln	Ser	Gln	Thr	Ile	Asn	Ser	Gln
				580					585					590		
55																

# EP 1 310 571 A2

Gly Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln  
595 600 605

5 Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro  
610 615 620

10 Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile  
625 630 635 640

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr  
645 650 655

15 Pro Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val  
660 665 670

20 Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp  
675 680 685

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val  
690 695 700

25 Glu Phe Ala Val Asn Pro Asp Gly Val Tyr Thr Glu Pro Arg Pro Ile  
705 710 715 720

30 Gly Thr Arg Tyr Leu Pro Arg Asn Leu  
725

<210> 110  
<211> 729  
<212> PRT  
35 <213> capsid protein of AAV serotype, clone F5VP1@3

<400> 110  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

40 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

45 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

50 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

55 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
85 90 95

# EP 1 310 571 A2

5 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 10 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 15 Pro Ile Asp Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln  
 145 150 155 160  
 Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu  
 165 170 175  
 20 Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser  
 180 185 190  
 25 Ser Val Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Thr Ala  
 195 200 205  
 Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp  
 210 215 220  
 30 His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr  
 225 230 235 240  
 35 Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile  
 245 250 255  
 Ser Ser Ser Ser Ser Gly Ala Thr Asn Asp Asn His Tyr Phe Gly Tyr  
 260 265 270  
 40 Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe  
 275 280 285  
 45 Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg  
 290 295 300  
 Pro Lys Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val  
 305 310 315 320  
 50 Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr  
 325 330 335  
 55 Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly  
 340 345 350

EP 1 310 571 A2

5 Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met  
355 360 365

Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val  
370 375 380

10 Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu  
385 390 395 400

15 Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Ser Phe Glu Asp Val Pro  
405 410 415

Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn  
420 425 430

20 Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr  
435 440 445

25 Thr Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr  
450 455 460

Met Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln  
465 470 475 480

30 Gln Arg Leu Ser Lys Asn Leu Asp Phe Asn Asn Asn Ser Asn Phe Ala  
485 490 495

35 Trp Thr Ala Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr  
500 505 510

Asn Pro Gly Ile Pro Met Ala Thr Asn Lys Asp Asp Glu Asp Gln Phe  
515 520 525

40 Phe Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn  
530 535 540

45 Lys Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys  
545 550 555 560

Thr Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn  
565 570 575

50 Leu Gln Ser Ser Thr Ala Gly Pro Gln Ser Gln Thr Ile Asn Ser Gln  
580 585 590

55 Gly Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln  
595 600 605

# EP 1 310 571 A2

5 Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro  
610 615 620

Ser Pro Leu Met Gly Gly Phe Gly Leu Glu His Pro Pro Pro Gln Ile  
625 630 635 640

10 Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr  
645 650 655

Pro Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val  
660 665 670

15 Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp  
675 680 685

20 Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val  
690 695 700

25 Glu Phe Ala Val Asn Pro Asp Gly Val Tyr Thr Glu Pro Arg Pro Ile  
705 710 715 720

Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
725

30 <210> 111  
<211> 729  
<212> PRT  
<213> capsid protein of AAV serotype, clone F3VP1

35 <400> 111  
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
1 5 10 15

40 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
35 40 45

45 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
50 55 60

50 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala  
85 90 95

55

# EP 1 310 571 A2

	Asp	Ala	Glu	Phe	Gln	Glu	Arg	Leu	Gln	Glu	Asp	Thr	Ser	Phe	Gly	Gly	
				100					105					110			
5	Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Val	Leu	Glu	Pro	
			115					120					125				
10	Leu	Gly	Leu	Val	Glu	Glu	Gly	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg	
		130					135					140					
	Pro	Ile	Gly	Ser	Pro	Asp	Ser	Ser	Thr	Gly	Ile	Gly	Lys	Lys	Gly	Gln	
	145					150					155					160	
15	Gln	Pro	Ala	Lys	Lys	Lys	Leu	Asn	Phe	Gly	Gln	Thr	Gly	Asp	Ser	Glu	
					165					170					175		
20	Ser	Val	Pro	Asp	Pro	Gln	Pro	Leu	Gly	Glu	Pro	Pro	Ala	Ala	Pro	Ser	
				180					185					190			
	Ser	Val	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly	Gly	Ala	Pro	Met	Ala	
			195					200					205				
25	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Asn	Ala	Ser	Gly	Asn	Trp	
		210					215					220					
30	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val	Ile	Thr	Thr	Ser	Thr	
	225					230					235					240	
	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	Leu	Tyr	Lys	Gln	Ile	
					245				250						255		
35	Ser	Ser	Ser	Ser	Ser	Gly	Ala	Thr	Asn	Asp	Asn	His	Tyr	Phe	Gly	Tyr	
					260				265					270			
40	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg	Phe	His	Cys	His	Phe	
			275					280					285				
	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn	Asn	Trp	Gly	Phe	Arg	
		290					295					300					
45	Pro	Lys	Lys	Leu	Arg	Phe	Lys	Leu	Leu	Asn	Ile	Gln	Val	Lys	Glu	Val	
	305					310					315					320	
50	Thr	Thr	Asn	Asp	Gly	Val	Thr	Thr	Ile	Ala	Asn	Asn	Leu	Thr	Ser	Thr	
					325					330					335		
	Val	Gln	Val	Phe	Ser	Asp	Ser	Glu	Tyr	Gln	Leu	Pro	Tyr	Val	Leu	Gly	
				340					345					350			
55																	

# EP 1 310 571 A2

Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met  
 355 360 365  
 5  
 Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asp Asn Gly Ser Gln Ser Val  
 370 375 380  
 10  
 Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu  
 385 390 395 400  
 Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Ser Phe Glu Asp Val Pro  
 405 410 415  
 15  
 Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn  
 420 425 430  
 20  
 Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr  
 435 440 445  
 Thr Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr  
 450 455 460  
 25  
 Met Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln  
 465 470 475 480  
 30  
 Gln Arg Leu Ser Lys Asn Leu Asp Phe Asn Asn Asn Ser Asn Phe Ala  
 485 490 495  
 Trp Thr Ala Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr  
 500 505 510  
 35  
 Asn Pro Gly Ile Pro Met Ala Thr Asn Lys Asp Asp Glu Asp Gln Phe  
 515 520 525  
 40  
 Phe Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn  
 530 535 540  
 Lys Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys  
 545 550 555 560  
 45  
 Thr Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn  
 565 570 575  
 50  
 Ile Gln Ser Ser Thr Ala Gly Pro Gln Ser Gln Thr Ile Asn Ser Gln  
 580 585 590  
 Gly Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln  
 595 600 605  
 55

# EP 1 310 571 A2

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro  
 610 615 620  
 5  
 Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile  
 625 630 635 640  
 10  
 Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr  
 645 650 655  
 Pro Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val  
 660 665 670  
 15  
 Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp  
 675 680 685  
 20  
 Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val  
 690 695 700  
 Glu Phe Ala Val Asn Pro Asp Gly Val Tyr Thr Glu Pro Arg Pro Ile  
 705 710 715 720  
 25  
 Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
 725  
 30  
 <210> 112  
 <211> 735  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone 42.6B  
 <400> 112  
 35  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
 40  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 35 40 45  
 45  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
 50  
 Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 85 90 95  
 55  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110

EP 1 310 571 A2

5 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125  
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg  
 130 135 140  
 10 Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile  
 145 150 155 160  
 Gly Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln  
 15 165 170 175  
 Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro  
 180 185 190  
 20 Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly  
 195 200 205  
 Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser  
 25 210 215 220  
 Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val  
 225 230 235 240  
 30 Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His  
 245 250 255  
 Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp  
 35 260 265 270  
 Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn  
 275 280 285  
 40 Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn  
 290 295 300  
 Asn Asn Trp Gly Phe Arg Pro Arg Lys Leu Arg Phe Lys Leu Phe Asn  
 45 305 310 315 320  
 Ile Gln Val Lys Glu Val Thr Thr Asp Asp Gly Val Thr Thr Ile Ala  
 325 330 335  
 50 Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Ser Asp Ser Glu Tyr Gln  
 340 345 350  
 Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe  
 55 355 360 365

# EP 1 310 571 A2

5	Pro	Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	370	375	380
	Asn	Gly	Ser	Gln	Ser	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	385	390	395
10	Phe	Pro	Ser	Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Glu	Phe	Ser	Tyr	405	410	415
	Thr	Phe	Glu	Glu	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	420	425	430
15	Leu	Asp	Arg	Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	435	440	445
20	Ala	Arg	Thr	Gln	Ser	Thr	Thr	Gly	Ser	Thr	Arg	Glu	Leu	Gln	Phe	His	450	455	460
	Gln	Ala	Gly	Pro	Asn	Thr	Met	Ala	Glu	Gln	Ser	Lys	Asn	Trp	Leu	Pro	465	470	475
25	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Leu	Ser	Lys	Asn	Ile	Asp	Ser	Asn	485	490	495
30	Asn	Asn	Ser	Asn	Phe	Ala	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	500	505	510
	Gly	Arg	Asn	Ser	Leu	Thr	Asn	Pro	Gly	Val	Ala	Met	Ala	Thr	Asn	Lys	515	520	525
35	Asp	Asp	Glu	Asp	Gln	Phe	Phe	Pro	Ile	Asn	Gly	Val	Leu	Val	Phe	Gly	530	535	540
40	Lys	Thr	Gly	Ala	Ala	Asn	Lys	Thr	Thr	Leu	Glu	Asn	Val	Leu	Met	Thr	545	550	555
	Ser	Glu	Glu	Glu	Ile	Lys	Thr	Thr	Asn	Pro	Val	Ala	Thr	Glu	Glu	Tyr	565	570	575
45	Gly	Val	Val	Ser	Ser	Asn	Leu	Gln	Ser	Ser	Thr	Ala	Gly	Pro	Gln	Thr	580	585	590
50	Gln	Thr	Val	Asn	Ser	Gln	Gly	Ala	Leu	Pro	Gly	Met	Val	Trp	Gln	Asn	595	600	605
	Arg	Asp	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	Pro	His	Thr	610	615	620
55																			

# EP 1 310 571 A2

5 Asp Gly Asn Phe His Pro Ser Pro Leu Met Asp Gly Phe Gly Leu Lys  
 625 630 635 640  
  
 His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn  
 645 650 655  
 10 Pro Pro Glu Val Phe Thr Pro Ala Lys Phe Ala Ser Phe Ile Thr Gln  
 660 665 670  
  
 Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys  
 15 675 680 685  
  
 Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr  
 690 695 700  
 20 Ala Lys Ser Asn Asn Val Glu Phe Ala Val Asn Asn Glu Gly Val Tyr  
 705 710 715 720  
  
 Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
 25 725 730 735  
  
 <210> 113  
 <211> 685  
 <212> PRT  
 30 <213> capsid protein of AAV serotype, clone 42.12  
  
 <400> 113  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser  
 1 5 10 15  
 35 Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro  
 20 25 30  
  
 Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro  
 40 35 40 45  
  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60  
 45 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80  
  
 Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 50 85 90 95  
  
 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly  
 100 105 110  
 55

# EP 1 310 571 A2

	Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Val	Leu	Glu	Pro
			115					120					125			
5	Leu	Gly	Leu	Val	Glu	Glu	Gly	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg
		130					135					140				
	Pro	Val	Glu	Pro	Ser	Pro	Gln	Arg	Ser	Pro	Asp	Ser	Ser	Thr	Gly	Ile
10	145					150					155					160
	Gly	Lys	Thr	Gly	Gln	Gln	Pro	Ala	Lys	Lys	Arg	Leu	Asn	Phe	Gly	Gln
					165					170					175	
15	Thr	Gly	Asp	Ser	Glu	Ser	Val	Pro	Asp	Pro	Gln	Pro	Ile	Gly	Glu	Pro
				180					185					190		
	Pro	Ala	Gly	Pro	Ser	Gly	Leu	Gly	Ser	Gly	Thr	Met	Ala	Ala	Gly	Gly
20			195					200					205			
	Gly	Ala	Pro	Met	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser
		210					215					220				
25	Ser	Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val
	225					230					235					240
	Ile	Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His
30					245					250					255	
	Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ser	Thr	Asn	Asp
				260					265					270		
35	Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn
			275					280					285			
	Arg	Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn
40		290					295					300				
	Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Asn	Phe	Lys	Leu	Phe	Asn
	305					310					315				320	
45	Ile	Gln	Val	Lys	Glu	Val	Thr	Gln	Asn	Glu	Gly	Thr	Lys	Thr	Ile	Ala
					325					330					335	
	Asn	Asn	Leu	Thr	Ser	Thr	Ile	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln
50				340					345					350		
	Leu	Pro	Tyr	Val	Leu	Gly	Ser	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe
			355					360					365			
55																

# EP 1 310 571 A2

Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn  
 370 375 380  
 5  
 Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr  
 385 390 395 400  
 10  
 Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr  
 405 410 415  
 Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser  
 420 425 430  
 15  
 Leu Asp Arg Leu Thr Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu  
 435 440 445  
 20  
 Ala Arg Thr Gln Ser Thr Thr Gly Ser Thr Arg Gly Leu Gln Phe His  
 450 455 460  
 Gln Ala Gly Pro Asn Thr Met Ala Glu Gln Ser Lys Asn Trp Leu Pro  
 465 470 475 480  
 25  
 Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Asn Ile Asp Ser Asn  
 485 490 495  
 30  
 Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu Asn  
 500 505 510  
 Gly Arg Asn Ser Leu Thr Asn Pro Gly Val Ala Met Ala Thr Asn Lys  
 515 520 525  
 35  
 Asp Asp Glu Asp Gln Phe Phe Pro Ile Asn Gly Val Leu Val Phe Gly  
 530 535 540  
 40  
 Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu Met Thr  
 545 550 555 560  
 Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Glu Tyr  
 565 570 575  
 45  
 Gly Val Val Ser Ser Asn Leu Gln Ser Ser Thr Ala Gly Pro Gln Thr  
 580 585 590  
 50  
 Gln Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val Trp Gln Asn  
 595 600 605  
 Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr  
 610 615 620  
 55

# EP 1 310 571 A2

Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys  
 625 630 635 640  
 5 His Pro Pro Pro Gln Ile Leu Ile Lys Tyr Thr Ser Asn Tyr Tyr Lys  
 645 650 655  
 10 Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu Gly Thr Tyr Ser Glu  
 660 665 670  
 Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu  
 675 680 685  
 15 <210> 114  
 <211> 724  
 <212> PRT  
 <213> capsid protein of AAV serotype, clone AAV5CAP  
 20 <400> 114  
 Met Ser Phe Val Asp His Pro Pro Asp Trp Leu Glu Glu Val Gly Glu  
 1 5 10 15  
 25 Gly Leu Arg Glu Phe Leu Gly Leu Glu Ala Gly Pro Pro Lys Pro Lys  
 20 25 30  
 Pro Asn Gln Gln His Gln Asp Gln Ala Arg Gly Leu Val Leu Pro Gly  
 35 40 45  
 30 Tyr Asn Tyr Leu Gly Pro Gly Asn Gly Leu Asp Arg Gly Glu Pro Val  
 50 55 60  
 35 Asn Arg Ala Asp Glu Val Ala Arg Glu His Asp Ile Ser Tyr Asn Glu  
 65 70 75 80  
 Gln Leu Glu Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp  
 85 90 95  
 40 Ala Glu Phe Gln Glu Lys Leu Ala Asp Asp Thr Ser Phe Gly Gly Asn  
 100 105 110  
 45 Leu Gly Lys Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro Phe  
 115 120 125  
 Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Thr Gly Lys Arg Ile  
 130 135 140  
 50 Asp Asp His Phe Pro Lys Arg Lys Lys Ala Arg Thr Glu Glu Asp Ser  
 145 150 155 160  
 55 Lys Pro Ser Thr Ser Ser Asp Ala Glu Ala Gly Pro Ser Gly Ser Gln  
 165 170 175

EP 1 310 571 A2

5 Gln Leu Gln Ile Pro Ala Gln Pro Ala Ser Ser Leu Gly Ala Asp Thr  
180 185 190

Met Ser Ala Gly Gly Gly Gly Pro Leu Gly Asp Asn Asn Gln Gly Ala  
195 200 205

10 Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys Asp Ser Thr Trp  
210 215 220

15 Met Gly Asp Arg Val Val Thr Lys Ser Thr Arg Thr Trp Val Leu Pro  
225 230 235 240

Ser Tyr Asn Asn His Gln Tyr Arg Glu Ile Lys Ser Gly Ser Val Asp  
245 250 255

20 Gly Ser Asn Ala Asn Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr  
260 265 270

25 Phe Asp Phe Asn Arg Phe His Ser His Trp Ser Pro Arg Asp Trp Gln  
275 280 285

Arg Leu Ile Asn Asn Tyr Trp Gly Phe Arg Pro Arg Ser Leu Arg Val  
290 295 300

30 Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Val Gln Asp Ser Thr  
305 310 315 320

35 Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp  
325 330 335

Asp Asp Tyr Gln Leu Pro Tyr Val Val Gly Asn Gly Thr Glu Gly Cys  
340 345 350

40 Leu Pro Ala Phe Pro Pro Gln Val Phe Thr Leu Pro Gln Tyr Gly Tyr  
355 360 365

45 Ala Thr Leu Asn Arg Asp Asn Thr Glu Asn Pro Thr Glu Arg Ser Ser  
370 375 380

Phe Phe Cys Leu Glu Tyr Phe Pro Ser Lys Met Leu Arg Thr Gly Asn  
385 390 395 400

50 Asn Phe Glu Phe Thr Tyr Asn Phe Glu Glu Val Pro Phe His Ser Ser  
405 410 415

55 Phe Ala Pro Ser Gln Asn Leu Phe Lys Leu Ala Asn Pro Leu Val Asp  
420 425 430

EP 1 310 571 A2

5	Gln Tyr Leu Tyr Arg Phe Val Ser Thr Asn Asn Thr Gly Gly Val Gln	435	440	445
	Phe Asn Lys Asn Leu Ala Gly Arg Tyr Ala Asn Thr Tyr Lys Asn Trp	450	455	460
10	Phe Pro Gly Pro Met Gly Arg Thr Gln Gly Trp Asn Leu Gly Ser Gly	465	470	475
				480
15	Val Asn Arg Ala Ser Val Ser Ala Phe Ala Thr Thr Asn Arg Met Glu	485	490	495
	Leu Glu Gly Ala Ser Tyr Gln Val Pro Pro Gln Pro Asn Gly Met Thr	500	505	510
20	Asn Asn Leu Gln Gly Ser Asn Thr Tyr Ala Leu Glu Asn Thr Met Ile	515	520	525
	Phe Asn Ser Gln Pro Ala Asn Pro Gly Thr Thr Ala Thr Tyr Leu Glu	530	535	540
25				
	Gly Asn Met Leu Ile Thr Ser Glu Ser Glu Thr Gln Pro Val Asn Arg	545	550	555
				560
30	Val Ala Tyr Asn Val Gly Gly Gln Met Ala Thr Asn Asn Gln Ser Ser	565	570	575
	Thr Thr Ala Pro Ala Thr Gly Thr Tyr Asn Leu Gln Glu Ile Val Pro	580	585	590
35				
	Gly Ser Val Trp Met Glu Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp	595	600	605
40				
	Ala Lys Ile Pro Glu Thr Gly Ala His Phe His Pro Ser Pro Ala Met	610	615	620
45	Gly Gly Phe Gly Leu Lys His Pro Pro Pro Met Met Leu Ile Lys Asn	625	630	635
				640
	Thr Pro Val Pro Gly Asn Ile Thr Ser Phe Ser Asp Val Pro Val Ser	645	650	655
50				
	Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Thr Val Glu Met Glu	660	665	670
55	Trp Glu Leu Lys Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln	675	680	685

# EP 1 310 571 A2

Tyr Thr Asn Asn Tyr Asn Asp Pro Gln Phe Val Asp Phe Ala Pro Asp  
 5 690 695 700

Ser Thr Gly Glu Tyr Arg Thr Thr Arg Pro Ile Gly Thr Arg Tyr Leu  
 705 710 715 720

10 Thr Arg Pro Leu

15 <210> 115  
 <211> 9  
 <212> DNA  
 <213> DraIII restriction enzyme site  
 <400> 115

20 caccacgtc 9

25 <210> 116  
 <211> 28  
 <212> DNA  
 <213> AV2cas  
 <400> 116

30 cgcagagacc aaagttcaac tgaaacga 28

35 <210> 117  
 <211> 255  
 <212> DNA  
 <213> adeno-associated virus serotype 10  
 <400> 117

40 ggtaattcct ccggaaattg gcattgcgat tccacatggc tgggcgacag agtcatcacc 60  
 accagcaccg gaacctgggt cctgcccacc tacaacaacc acatctacaa gcaaattctcc 120  
 agcgagacag gagccaccaa cgacaaccac tacttcggct acagcaccgg ctgggggtat 180  
 tttgacttta acagattcca ctgccacttt tcaccacgtg actggcagcg actcatcaac 240  
 aacaactggg gatttc 255

45 <210> 118  
 <211> 258  
 <212> DNA  
 <213> adeno-associated virus serotype 11  
 <400> 118

50 ggtaattcct ccggaaattg gcattgcgat tccacatggc tgggcgacag agtcatcacc 60  
 accagcaccg gaacctgggc cctgccaacc tacaacaacc acctctacaa acaaattctcc 120  
 55 agcgcttcaa cgggggcccag caacgacaac cactactttg gctacagcac cccctggggg 180

# EP 1 310 571 A2

tatatttgact ttaacagatt ccactgccac ttctcaccac gtgactggca ggcactcatc 240  
5 aacaacaact ggggattc 258

<210> 119  
<211> 255  
10 <212> DNA  
<213> adeno-associated virus serotype 12

<400> 119  
ggtaattcct ccggaattg gcattgcgat tccacatggc tgggcgaccg agtcattacc 60  
15 accagcacc cggacttgggc cctgccacc tacaacaacc acctctacaa gcaaattctc 120  
agccaatcgg gtgccaccaa cgacaaccac tacttcggct acagcacc cttgggggtat 180  
tttgatttca acagattcca ctgccatttc tcaccacgtg actggcagcg actcatcaac 240  
20 aacaactggg gattc 255

<210> 120  
<211> 2205  
25 <212> DNA  
<213> adeno-associated virus serotype, clone A3.1vp1

<400> 120  
atggctgccg atggttatct tccagattgg ctgcaggaca ctctctctga aggaatcaga 60  
30 cagtgggtgga agctcaaacc tggccacca ccgccgaaac ctaaccaaca acaccgggac 120  
gacagtaggg gtcttgtgct tcctgggtac aagtacctcg gacccttcaa cggactcgac 180  
aaaggagagc cgttcaacga ggcagacgcc gcggccctcg agcacgacaa agcctacgac 240  
35 caccagctca agcaagggga caaccctgac ctcaaataca accacgcgga cgtgaattc 300  
caggagcgtc ttcaagaaga tacgtctttc gggggcaacc tcgggcgagc agtcttccag 360  
gccaaaaaga ggtactcga gcctcttggt ctggttgagg aagctgttaa gacggctctc 420  
40 ggaaaaaaga gacctataga gcagtctcct gcagaaccgg actcttctc gggcatcggc 480  
aatcaggcc agcagccgc taagaaaaga ctcaattttg gtcagactgg cgacacagag 540  
tcagtccag acctcaacc aatcggagaa cccccgcag cccctcttg tgtgggatct 600  
45 aatacaatgg cttcaggcgg tggggacca atggcagaca ataacgaagg cggcgacgga 660  
gtgggtaatt cctcgggaaa ttggcattgc gattccacat ggatgggga cagagttatc 720  
accaccagca caagaacctg ggccctcccc acctacaata atcacctcta caagcaaadc 780  
50 tccagcgaat cgggagccac caacgacaac cactacttcg gctacagcac ccctggggg 840  
tatatttgact ttaacagatt ccactgtcac ttctcaccac gtgactggca ggcactcatc 900  
aacaacaact ggggatttag acccaagaaa ctcaatttca agctcttcaa catccaagtc 960  
55 aaggagggtca cgcagaatga tggaaccacg accatcgcca ataaccttac cagcacggtg 1020

	cagggtcttca cagactctga gtaccagctg ccctacgtcc tcgggttcggc tcaccagggc	1080
5	tgccttccgc cgttcccagc agacgtcttc atgattcctc agtacggcta cttgactctg	1140
	aacaatggca gccaaagcgt aggacgttct tcattctact gtctagagta ttttccctct	1200
	cagatgctga ggacgggaaa caacttcacc ttcagctaca cttttgaaga cgtgcctttc	1260
10	cacagcagct acgcgcacag ccagagtctg gatcggctga tgaatcctct cattgaccag	1320
	tacctgtatt acctgagcaa aactcagggg acaagtggaa caacgcagca atcgagactg	1380
	cagttcagcc aagctggggc tagctccatg gctcagcagg ccaaaaactg gctaccggga	1440
15	cccagctacc gacagcagcg aatgtctaag acggctaata acaacaacaa cagtgaattt	1500
	gcttggactg cagccaccaa atattacctg aatggaagaa attctctggg caatcccggg	1560
	cccccaatgg ccagtcacaa ggacgatgag gaaaagtatt tccccatgca cggaaatctc	1620
20	atctttggaa aacaaggcac aggaactacc aatgtggaca ttgaatcagt gcttattaca	1680
	gacgaagaag aaatcagaac aactaatcct gtggctacag aacaatacgg acaggttgcc	1740
	accaaccatc agagtcagaa caccacagct tcctatggaa gtgtggacag ccagggaatc	1800
25	ttacctggaa tgggtgtggc ggaccgcgat gtctatcttc aagggtcccat ttgggccaaa	1860
	actctcaca cggacggaca ctttcacctt tctccgctca tgggaggctt tggactgaaa	1920
	cacctcctc cccagatcct gatcaaaaac acacctgtgc cagcgaatcc cgcgaccact	1980
30	ttcactcctg gaaagtttgc ttcgttcatt acccagtatt ccaccggaca ggtcagcgtg	2040
	gaaatagagt gggagctgca gaaagaaaac agcaaacgct ggaaccaga aattcagtac	2100
	acctccaact acaacaagtc ggtgaatgtg gagtttaccg tggacgcaaa cgggtgttat	2160
35	tctgaacccc gccctattgg cactcgttac cttaccggga acttg	2205

#### 40 Claims

1. A method of detecting adeno-associated virus (AAV) sequences in a sample, said method comprising the steps of:

- 45 (a) subjecting a sample containing DNA to amplification via polymerase chain reaction (PCR) using a first set of primers which specifically amplify a first region of AAV nucleic acid sequences;
- (b) optionally subjecting the DNA to further amplification using a second set of primers which specifically amplify a second region which comprises the first region of AAV sequences and sequences which are 5' to the first region, such that AAV 5' extension sequences which anneal to the 5' end of the AAV sequences amplified by the primers for the first region are obtained;
- 50 (c) optionally subjecting the DNA to further amplification using a third set of primers which specifically amplify a third region which comprises the first region of AAV sequences and sequences which are 3' to the first region, such that AAV 3' extension sequences which anneal to the 3' end of the AAV sequences amplified by the primers for the first region are obtained,

55 wherein each of said regions is predetermined based upon the alignment of the nucleic acid sequences of at least two AAV serotypes and each of said regions comprises nucleic acid sequences which are highly conserved over at least 18 base pairs at the 5' end, optionally variable sequences in the middle, and sequences which are highly conserved over at least 18 base pairs at the 3' end of the sequences of the region, relative to the sequences

of the at least two aligned AAV serotypes; and

wherein each of the sets of primers consist of a 5' primer and a 3' primer;

wherein the presence of amplified sequences indicates the presence of an AAV in the sample.

- 5     **2.** A method according to claim 1, wherein steps (b) and (c) are performed, and further comprising the step of:  
  
          (d) using the amplified sequences to construct a sequence comprising a partial and/or complete AAV gene, thereby isolating an AAV gene sequence from the sample.
- 10    **3.** The method according to claim 2, wherein the first region is in the AAV capsid, the second region extends to the 3' end of the rep genes, and the third region extends to 5' of the AAV 3' ITR, such that step (d) permits the construction of AAV gene sequences comprising the complete AAV capsid gene.
- 4.** The method according to claim 2, further comprising the step of subjecting the DNA to amplification with a fourth  
15    set of primers which specifically amplifies a fourth region which extend 5' to the third region.
- 5.** The method according to claim 4, wherein the fourth region extends to 3' of the AAV 5' ITR, such that step (d) permits the construction of AAV gene sequences comprising the complete AAV rep and cap genes.
- 20    **6.** The method according to claim 1 or claim 2, wherein said sample comprises cellular or genomic DNA.
- 7.** The method according to claim 1 or claim 2, wherein said DNA has been extracted from the group consisting of cells, cell culture, tissue, tissue culture, and biological fluids.
- 25    **8.** The method according to claim 1 or claim 2, wherein the first region is about 250 base pairs in length.
- 9.** The method according to claim 1 or 2, wherein the first region is highly conserved over at least about 25 base pairs at the 5' and/or 3' end of the region.
- 30    **10.** The method according to claim 1 or 2, wherein the first region is highly conserved over at least about 30 base pairs at the 5' and/or 3' end of the region.
- 11.** The method according to claim 1 or claim 2, wherein the highly conserved sequences of the first region have at least 80% identity among the aligned AAV serotypes at the 5' and/or 3' end of the region.
- 35    **12.** The method according claim 11, wherein the highly conserved sequences of the first region have at least 90% identity among the aligned AAV serotypes at the 5' and/or 3' end of the region.
- 13.** The method according to claim 1 or claim 2, wherein the variable sequences in the middle of the first region have less than 70% identity among the aligned AAV serotypes.
- 40    **14.** The method according to claim 1 or claim 2, wherein the first region spans about bp 2800 to about 3200 of AAV 1, SEQ ID NO:6, and corresponding base pairs in other AAV serotypes.
- 45    **15.** The method according to claim 14, wherein the first region is 257 bp spanning bp 2886 to about 3143 bp of AAV 1, SEQ ID NO:6, and corresponding base pairs in other AAV serotypes.
- 16.** The method according to claim 1, wherein the primers are AV1ns and AV2cas.
- 50    **17.** A method of identifying the serotype of adeno-associated virus (AAV) sequences in a sample as known or unknown, said method comprising the steps of:  
  
          (a) obtaining an enzymatic digestion analysis of a sample containing DNA molecules comprising AAV sequences which span all or a portion of bp 2886 to 3143 of AAV1, SEQ ID NO:1, and corresponding regions  
55    of other AAV serotypes; and  
          (b) comparing the enzymatic digestion analysis from the sample to enzymatic digestion analysis for corresponding regions of one or more AAV serotypes, thereby identifying the AAV sequences in the sample as being from one of the one or more AAV serotypes or from an unknown serotype.

18. The method according to claim 17, wherein the AAV sequences were obtained by the method of claim 1.

19. A diagnostic kit for detecting the presence of an unknown adeno-associated virus (AAV) in a sample, said kit comprising:

- (a) a first set of primers which specifically amplify a first region of AAV nucleic acid sequences;
- (b) optionally a second set of primers specific for a second region of the AAV nucleic acid sequences which comprises the first region of AAV sequences and sequences which are 5' to the first region, such that AAV 5' extension sequences which anneal to the 5' end of the AAV sequences amplified by the primers for the first region are obtained;
- (c) optionally a third set of primers which specifically amplify a third region which comprises the first region of AAV sequences and sequences which are 3' to the first region, such that AAV 3' extension sequences which anneal to the 3' end of the AAV sequences amplified by the primers for the first region are obtained,

wherein each of said regions is predetermined based upon the alignment of the nucleic acid sequences of at least two AAV serotypes and each of said regions comprises nucleic acid sequences which are highly conserved over at least 18 base pairs at the 5' end, optionally variable sequences in the middle, and sequences which are highly conserved over at least 18 base pairs at the 3' end of the sequences of the region, relative to the sequences of the at least two aligned AAV serotypes

wherein each of the sets of primers consist of a 5' primer and a 3' primer.

20. A method for isolating novel adeno-associated viruses (AAV) from a cell, said method comprising the steps of: (a) infecting the cell with a virus which provides helper functions to said AAV; (b) isolating infectious clones containing AAV; (c) sequencing the isolated AAV; and (d) comparing the sequences of the isolated AAV to known AAV serotypes, whereby differences in the sequences of the isolated AAV and known AAV serotypes indicates the presence of a novel AAV.

21. The method according to claim 20, wherein said virus is an adenovirus.

22. The method according to claim 21, wherein said adenovirus is of human or non-human primate origin.

23. A novel adeno-associated virus (AAV) serotype identified by a method according to claim 2 or claim 20.

24. An isolated adeno-associated virus (AAV) comprising an AAV capsid having an amino acid sequence selected from the group consisting of:

AAV7, amino acids 1 to 737 of SEQ ID NO:2; C1, SEQ ID NO:60; C2, SEQ ID NO:61; C5, SEQ ID NO:62; A3-3, SEQ ID NO:66; A3-7, SEQ ID NO:67; A3-4, SEQ ID NO:68; A3-5, SEQ ID NO: 69; 3.3b, SEQ ID NO: 62; 223.4, SEQ ID NO: 73; 223-5, SEQ ID NO:74; 223-10, SEQ ID NO:75; 223-2, SEQ ID NO:76; 223-7, SEQ ID NO: 77; 223-6, SEQ ID NO: 78; 44-1, SEQ ID NO: 79; 44-5, SEQ ID NO:80; 44-2, SEQ ID NO:81; 42-15, SEQ ID NO: 84; 42-8, SEQ ID NO: 85; 42-13, SEQ ID NO:86; 42-3A, SEQ ID NO:87; 42-4, SEQ ID NO:88; 42-5A, SEQ ID NO:89; 42-1B, SEQ ID NO:90; 42-5B, SEQ ID NO:91; 43-1, SEQ ID NO: 92; 43-12, SEQ ID NO: 93; 43-5, SEQ ID NO:94; 43-21, SEQ ID NO:96; 43-25, SEQ ID NO: 97; 43-20, SEQ ID NO:99; 24.1, SEQ ID NO: 101; 42.2, SEQ ID NO:102; 7.2, SEQ ID NO: 103; 27.3, SEQ ID NO: 104; 16.3, SEQ ID NO: 105; 42.10, SEQ ID NO: 106; 42-3B, SEQ ID NO: 107; 42-11, SEQ ID NO: 108; F1, SEQ ID NO: 109; F5, SEQ ID NO: 110; F3, SEQ ID NO:111; 42-6B, SEQ ID NO: 112; and 42-12, SEQ ID NO: 113.

25. A protein comprising a fragment of an AAV capsid protein, said fragment selected from the group consisting of:

vp2 capsid protein, amino acids (aa) 138 to 737;  
vp3 capsid protein, aa 203 to 737;  
hypervariable region (HVR)1 through 12: aa 146 to 152; aa 182 to 187; aa 262 to 264; aa 263 to 266; aa 263 to 266; aa 381 to 383; aa 383 to 385; aa 450 to 474; aa 451 to 475; aa 490 to 495; aa 491 to 496; aa500 to 504; aa 501 to 505; aa 514 to 522; aa 533 to 554; aa 534 to 555; aa 581 to 594; aa 583 to 596; aa 658 to 667; aa 660 to 669; and aa 705 to 719; aa 707 to 772;  
aa 24-42, aa 25-28; aa 81-85; aa133 to 165; aa 134 - 165; aa 137 to 143; aa 154 to 156; aa 194 to 208; aa 261 to 274; aa 262 to 274; aa 171 to 173; aa 413 to 417; aa 449 to 478; as 494 to 525; as 534 to 571; aa 581 to 601; as 660 to 671; as 709 to 723; and

aa 1 to 184, aa 199 to 259; aa 274 to 446; aa 603 to 659; aa 670 to 706; aa 724 to 736; aa 185 to 198; aa 260 to 273; aa 447 to 477; aa 495 to 602; aa 660 to 669; and aa 707 to 723,

wherein the amino acid numbers are those of the AAV7 capsid, SEQ ID NO:2, and corresponding regions in the capsid of C1, SEQ ID NO:60; C2, SEQ ID NO:61; C5, SEQ ID NO:62; A3-3, SEQ ID NO:66; A3-7, SEQ ID NO:67; A3-4, SEQ ID NO:68; A3-5, SEQ ID NO: 69; 3.3b, SEQ ID NO: 62; 223.4, SEQ ID NO: 73; 223-5, SEQ ID NO:74; 223-10, SEQ ID NO:75; 223-2, SEQ ID NO:76; 223-7, SEQ ID NO: 77; 223-6, SEQ ID NO: 78; 44-1, SEQ ID NO: 79; 44-5, SEQ ID NO:80; 44-2, SEQ ID NO:81; 42-15, SEQ ID NO: 84; 42-8, SEQ ID NO: 85; 42-13, SEQ ID NO:86; 42-3A, SEQ ID NO:87; 42-4, SEQ ID NO:88; 42-5A, SEQ ID NO:89; 42-1B, SEQ ID NO:90; 42-5B, SEQ ID NO:91; 43-1, SEQ ID NO: 92; 43-12, SEQ ID NO: 93; 43-5, SEQ ID NO:94; 43-21, SEQ ID NO:96; 43-25, SEQ ID NO: 97; 43-20, SEQ ID NO:99; 24.1, SEQ ID NO: 101; 42.2, SEQ ID NO:102; 7.2, SEQ ID NO: 103; 27.3, SEQ ID NO: 104; 16.3, SEQ ID NO: 105; 42.10, SEQ ID NO: 106; 42-3B, SEQ ID NO: 107; 42-11, SEQ ID NO: 108; F1, SEQ ID NO: 109; F5, SEQ ID NO: 110; F3, SEQ ID NO:111; 42-6B, SEQ ID NO: 112; and 42-12, SEQ ID NO: 113.

26. An artificial adeno-associated virus (AAV) capsid protein comprising one or more of the protein fragments according to claim 25.

27. A recombinant adeno-associated virus (AAV) comprising an artificial capsid according to claim 26.

28. A molecule comprising a nucleic acid sequence encoding a protein according to claim 25 to 26.

29. A molecule comprising a nucleic acid sequence encoding a novel adeno-associated virus (AAV) serotype capsid protein having an amino acid sequence selected from the group consisting of: AAV7, amino acids 1 to 737 of SEQ ID NO:2; C1, SEQ ID NO:60; C2, SEQ ID NO:61; C5, SEQ ID NO:62; A3-3, SEQ ID NO:66; A3-7, SEQ ID NO: 67; A3-4, SEQ ID NO:68; A3-5, SEQ ID NO: 69; 3.3b, SEQ ID NO: 62; 223.4, SEQ ID NO: 73; 223-5, SEQ ID NO: 74; 223-10, SEQ ID NO:75; 223-2, SEQ ID NO:76; 223-7, SEQ ID NO: 77; 223-6, SEQ ID NO: 78; 44-1, SEQ ID NO: 79; 44-5, SEQ ID NO:80; 44-2, SEQ ID NO:81; 42-15, SEQ ID NO: 84; 42-8, SEQ ID NO: 85; 42-13, SEQ ID NO:86; 42-3A, SEQ ID NO:87; 42-4, SEQ ID NO:88; 42-5A, SEQ ID NO:89; 42-1 B, SEQ ID NO:90; 42-5B, SEQ ID NO:91; 43-1, SEQ ID NO: 92; 43-12, SEQ ID NO: 93; 43-5, SEQ ID NO:94; 43-21, SEQ ID NO:96; 43-25, SEQ ID NO: 97; 43-20, SEQ ID NO:99; 24.1, SEQ ID NO: 101; 42.2, SEQ ID NO:102; 7.2, SEQ ID NO: 103; 27.3, SEQ ID NO: 104; 16.3, SEQ ID NO: 105; 42.10, SEQ ID NO: 106; 42-3B, SEQ ID NO: 107; 42-11, SEQ ID NO: 108; F1, SEQ ID NO: 109; F5, SEQ ID NO: 110; F3, SEQ ID NO:111; 42-6B, SEQ ID NO: 112; and 42-12, SEQ ID NO: 113.

30. A molecule comprising a nucleic acid sequence encoding a novel adeno-associated virus (AAV) serotype capsid protein, said nucleic acid sequence selected from the group consisting of: AAV5, SEQ ID NO:2; 42-2, SEQ ID NO:9; 42-8, SEQ ID NO:27; 42-15, SEQ ID NO:28; 42-5b, SEQ ID NO: 29; 42-1b, SEQ ID NO:30; 42-13, SEQ ID NO: 31; 42-3a, SEQ ID NO: 32; 42-4, SEQ ID NO:33; 42-5a, SEQ ID NO: 34; 42-10, SEQ ID NO:35; 42-3b, SEQ ID NO: 36; 42-11, SEQ ID NO: 37; 42-6b, SEQ ID NO:38; 43-1, SEQ ID NO: 39; 43-5, SEQ ID NO: 40; 43-12, SEQ ID NO:41; 43-20, SEQ ID NO:42; 43-21, SEQ ID NO: 43; 43-23, SEQ ID NO:44; 43-25, SEQ ID NO: 45; 44.1, SEQ ID NO:47; 44.5, SEQ ID NO:47; 223.10, SEQ ID NO:48; 223.2, SEQ ID NO:49; 223.4, SEQ ID NO:50; 223.5, SEQ ID NO: 51; 223.6, SEQ ID NO: 52; 223.7, SEQ ID NO: 53; A3.4, SEQ ID NO: 54; A3.5, SEQ ID NO:55; A3.7, SEQ ID NO: 56; A3.3, SEQ ID NO:57; 42.12, SEQ ID NO: 58; 44.2, SEQ ID NO: 59; AAV10, SEQ ID NO: 117; AAV11, SEQ ID NO: 118; AAV12, SEQ ID NO:119; A3.1, SEQ ID NO:120; and H6, SEQ ID NO: 25.

31. A molecule comprising a nucleic acid sequence encoding a fragment of an adeno-associated virus capsid protein, said nucleic acid sequence selected from the group consisting of:

vp1, nt 825 to 3049;  
vp2, nt 1234 to 3049;  
vp 3, nt 1434 to 3049;  
nt 468 to 3090; and  
nt 725 to 3090,

wherein the nucleotides numbers are of AAV7, SEQ ID NO:1 and correspond to sequences in 42-2, SEQ ID NO:9; 42-8, SEQ ID NO:27; 42-15, SEQ ID NO:28; 42-5b, SEQ ID NO: 29; 42-1b, SEQ ID NO:30; 42-13, SEQ ID NO: 31; 42-3a, SEQ ID NO: 32; 42-4, SEQ ID NO:33; 42-5a, SEQ ID NO: 34; 42-10, SEQ ID NO:35; 42-3b,

## EP 1 310 571 A2

SEQ ID NO: 36; 42-11, SEQ ID NO: 37; 42-6b, SEQ ID NO:38; 43-1, SEQ ID NO: 39; 43-5, SEQ ID NO: 40; 43-12 ,  
SEQ ID NO:41; 43-20, SEQ ID NO:42; 43-21, SEQ ID NO: 43; 43-23, SEQ ID NO:44; 43-25 ,SEQ ID NO: 45; 44.1,  
SEQ ID NO:47; 44.5, SEQ ID NO:47; 223.10, SEQ ID NO:48; 223.2 , SEQ ID NO:49; 223.4, SEQ ID NO:50; 223.5,  
SEQ ID NO: 51; 223.6, SEQ ID NO: 52; 223.7, SEQ ID NO: 53; A3.4, SEQ ID NO: 54; A3.5, SEQ ID NO:55; A3.7 ,  
SEQ ID NO: 56; A3.3 ,SEQ ID NO:57; 42.12, SEQ ID NO: 58; 44.2, SEQ ID NO: 59; AAV10, SEQ ID NO: 117;  
AAV11, SEQ ID NO: 118; AAV12 ,SEQ ID NO:119; A3.1, SEQ ID NO:120; and H6, SEQ ID NO: 25.

32. The molecule according to claim 26 or claim 28 to 31, wherein said molecule is a plasmid.

33. The molecule according to claim 26 or claim 28 to 31, wherein said molecule further comprises a functional AAV *rep* gene.

34. The molecule according to claim 30, wherein said nucleic acid sequence is the AAV7 sequence, SEQ ID NO:1.

35. A method of generating a recombinant adeno-associated virus (AAV) comprising an AAV serotype capsid comprising the steps of culturing a host cell containing: (a) a molecule according to any of claims 26 or claim 28 to 31 which encodes an adeno-associated virus capsid; (b) a functional rep gene; (c) a minigene comprising AAV inverted terminal repeats (ITRs) and a transgene; and (d) sufficient helper functions to permit packaging of the minigene into the AAV capsid protein.

36. A host cell transfected with an adeno-associated virus according to claim 24 or a molecule according to any of claims 26, or claims 28 to 31.

37. A composition comprising an AAV according to claim 24 or claim 27, and a physiologically compatible carrier.

38. A composition comprising a molecule according to any of claims 26, or claims 28 to 31 and a physiologically compatible carrier.

39. A method of delivering a transgene to a cell, said method comprising the step of contacting the cell with an AAV according to claim 24 or claim 27, wherein said rAAV comprises the transgene.

40. A molecule comprising a heterologous adeno-associated virus (AAV) serotype 7 nucleic acid sequence, said sequence comprising:

nucleotides (nt) 1 to 107 of SEQ ID NO: 1;  
nt 107 to 2215 of SEQ ID NO:1;  
nt 334 to 2215 of SEQ ID NO:1;  
nt 2222 to 4435 of SEQ ID NO:1.  
nt 2633 to 4435 of SEQ ID NO:1;  
nt 2831 to 4435 of SEQ ID NO:1; and  
nt 4704 to 4721 of SEQ ID NO: 1.

41. A molecule encoding an adeno-associated virus (AAV) serotype 7 rep protein or a fragment thereof, said protein or fragment selected from the group consisting of:

amino acid (aa) 1 to 623, aa 1 to 171; aa 172 to 372, aa 373 to 444, and aa 445 to 623 of SEQ ID NO:3.

42. A host cell containing a molecule according to claim 40 or 41.

FIG. 1A

	1				50
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	TTGCCCCTC	CCTCTCTGCG	CGCTCGCTCG	CTCGGTGGGG	CCTGCGGACC
AAV2	TTGCCCCTC	CCTCTCTGCG	CGCTCGCTCG	CTCACTGAGG	CCGGGCGACC
AAV3	TTGCCCCTC	CCTCTATGCG	CACTCGCTCG	CTCGGTGGGG	CCTGGCGACC
AAV8	.....	.....	.....	.....	.....
AAV9	.....	.....	.....	.....	.....
AAV7	TTGCCCCTC	CCTCTATGCG	CGCTCGCTCG	CTCGGTGGGG	CCTGCGGACC
44_2	.....	.....	.....	.....	.....

Fig. 1B

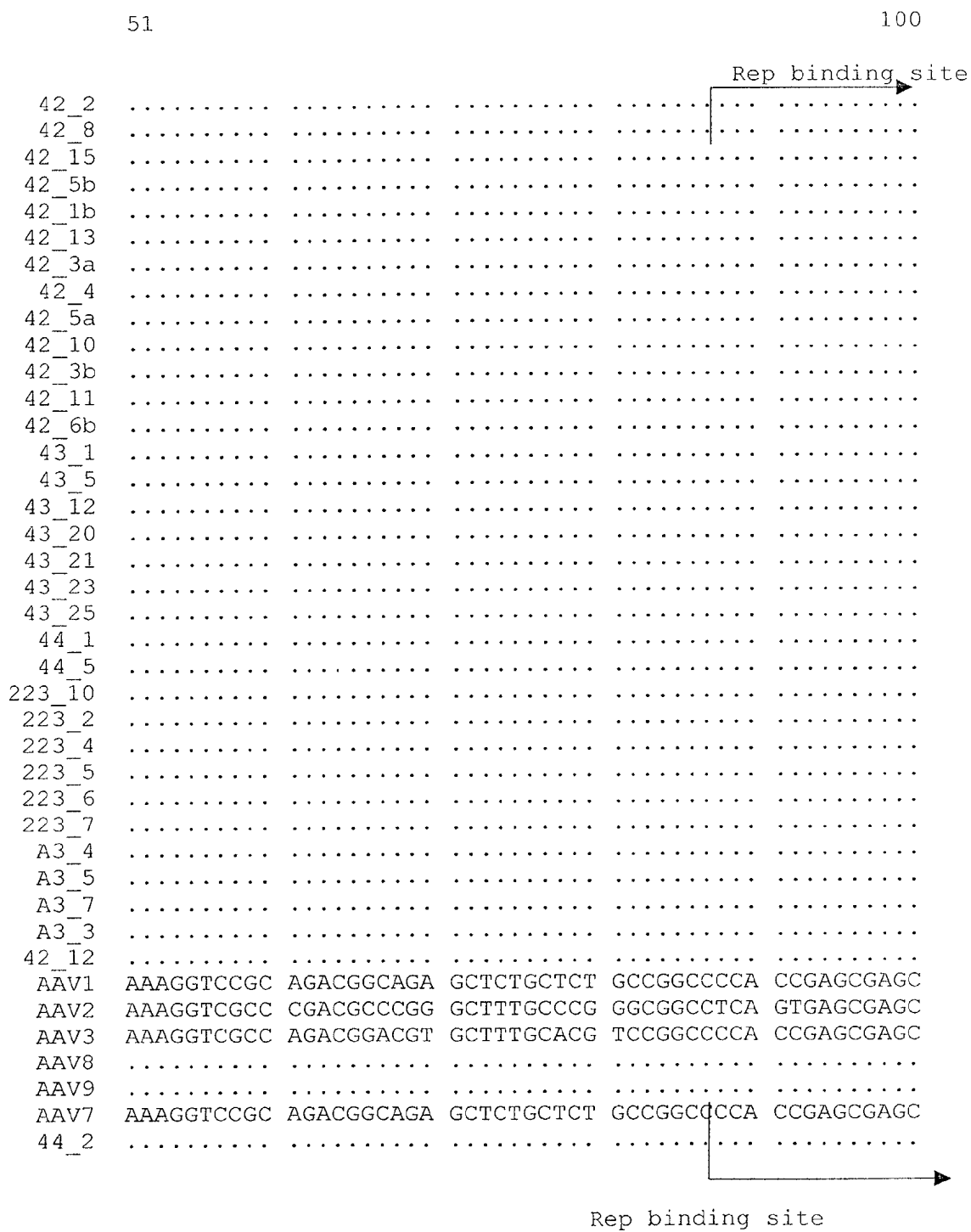


Fig. 1C

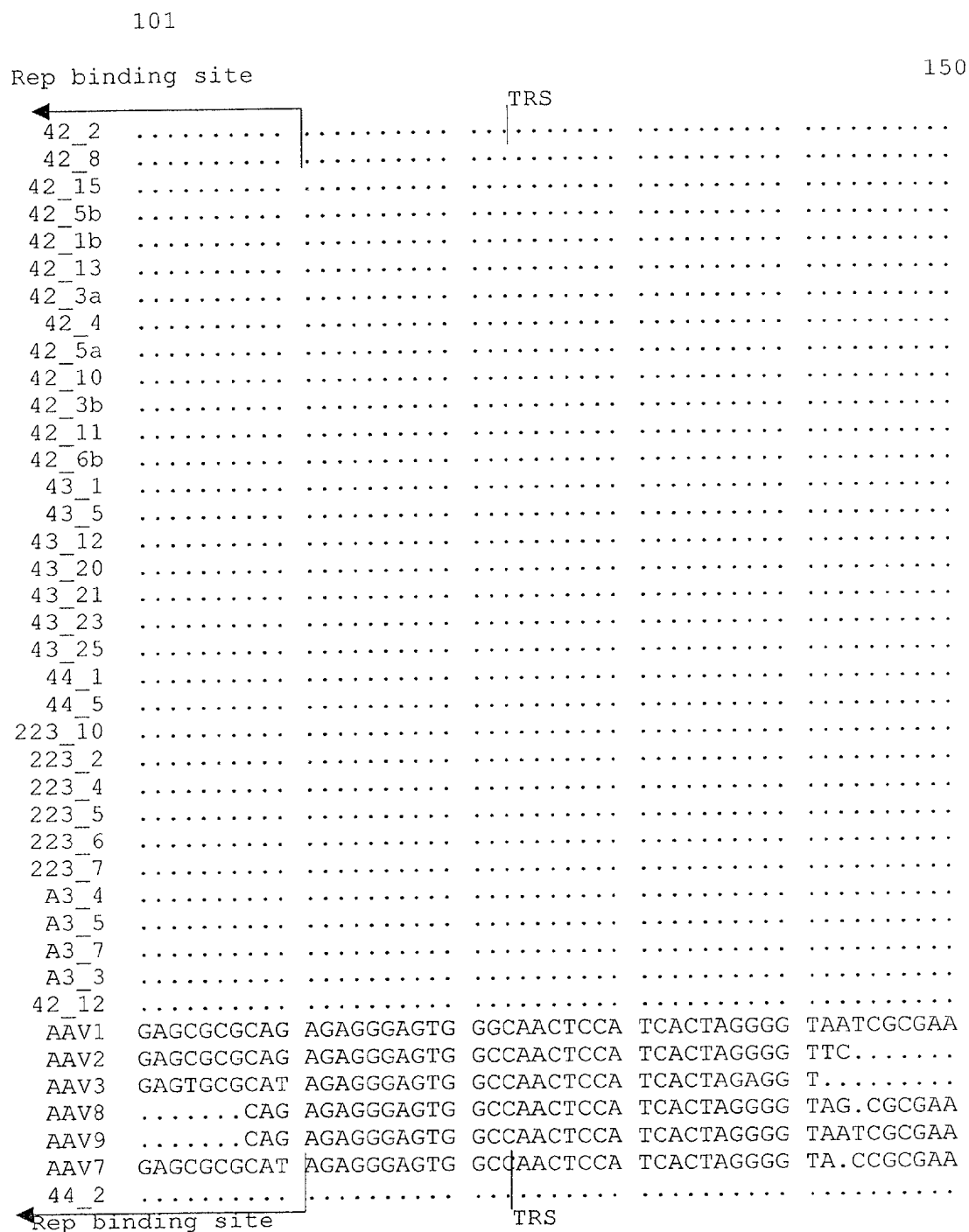


Fig. 1D

	151				200
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	GCGCCTCCCA	CGCTGCCGCG	TCAGCGCTGA	CGTAAATTAC	GTCATAGGGG
AAV2	.....CTG	GAGGGGTGGA	GTCGTGACGT	GAATTACGTC	ATAGGGTTAG
AAV3	.....ATG	GCAGTGACGT	AACGCGAAGC	GCGCGAAGCG	AGACCACGCC
AAV8	GCGCCTCCCA	CGCTGCCGCG	TCAGCGCTGA	CGTAAATTAC	GTCATAGGGG
AAV9	GCGCCTCCCA	CGCTGCCGCG	TCAGCGCTGA	CGTAGATTAC	GTCATAGGGG
AAV7	GCGCCTCCCA	CGCTGCCGCG	TCAGCGCTGA	CGTAAATCAC	GTCATAGGGG
44_2	.....	.....	.....	.....	.....

Fig. 1E

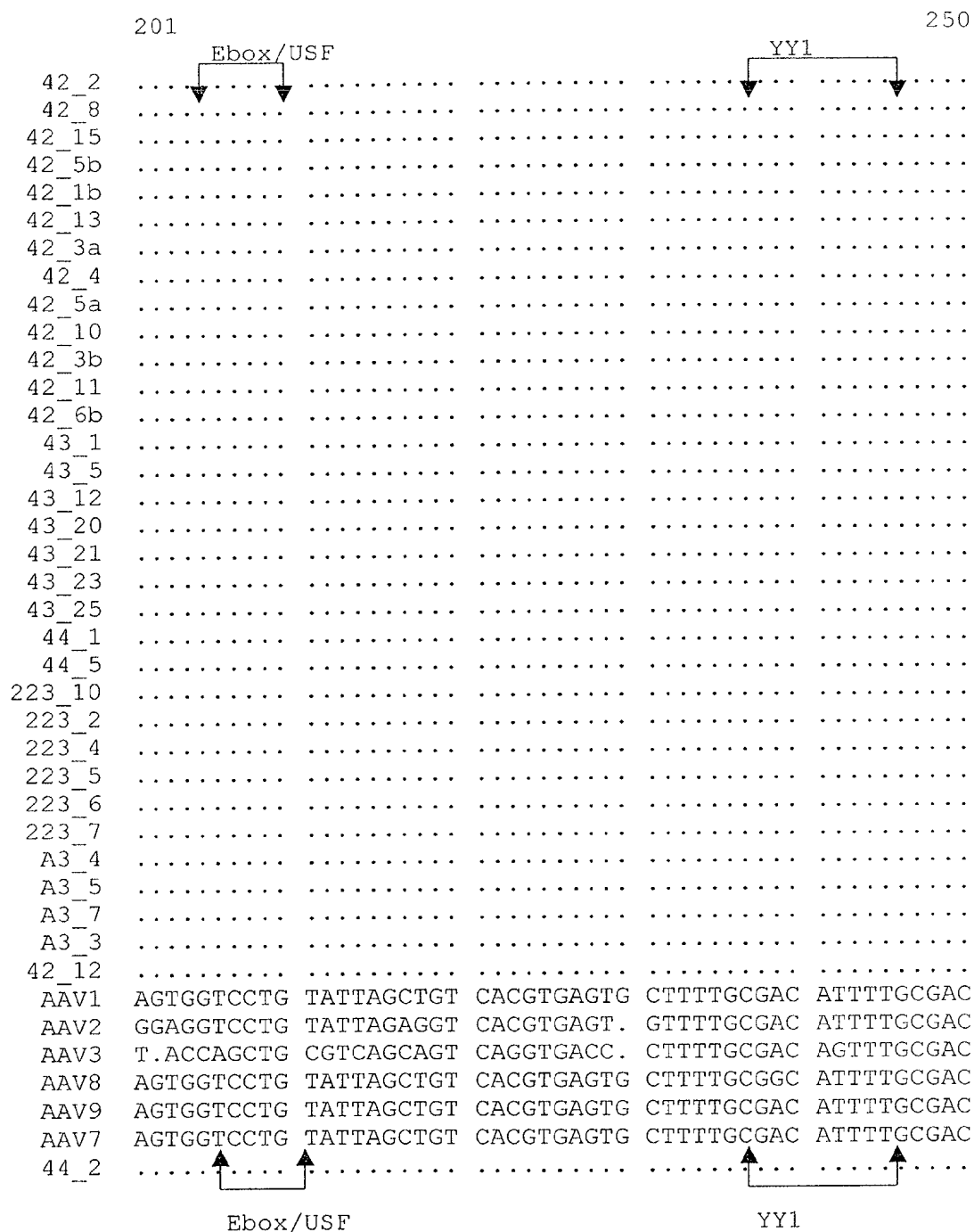


Fig. 1F

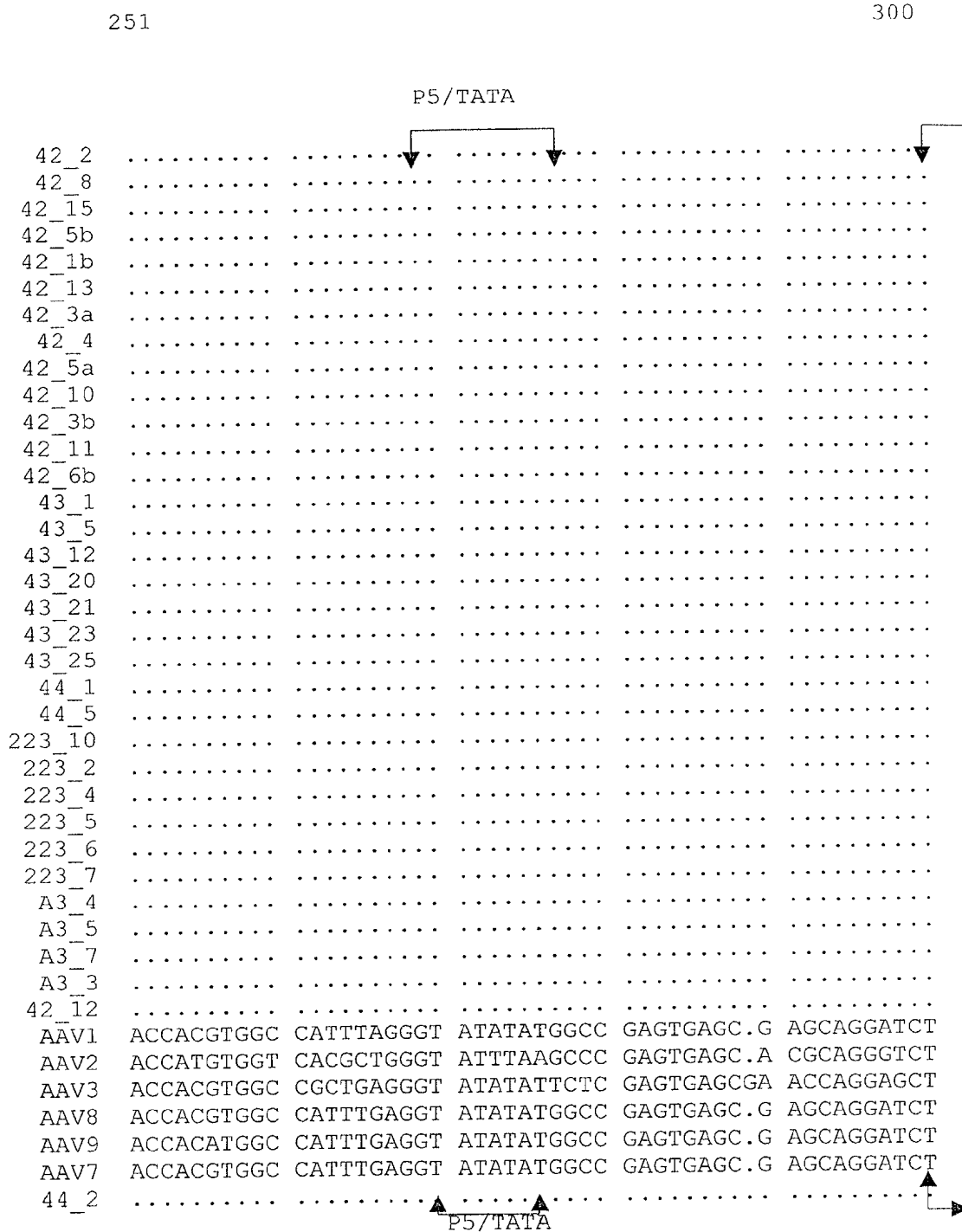


Fig. 1G

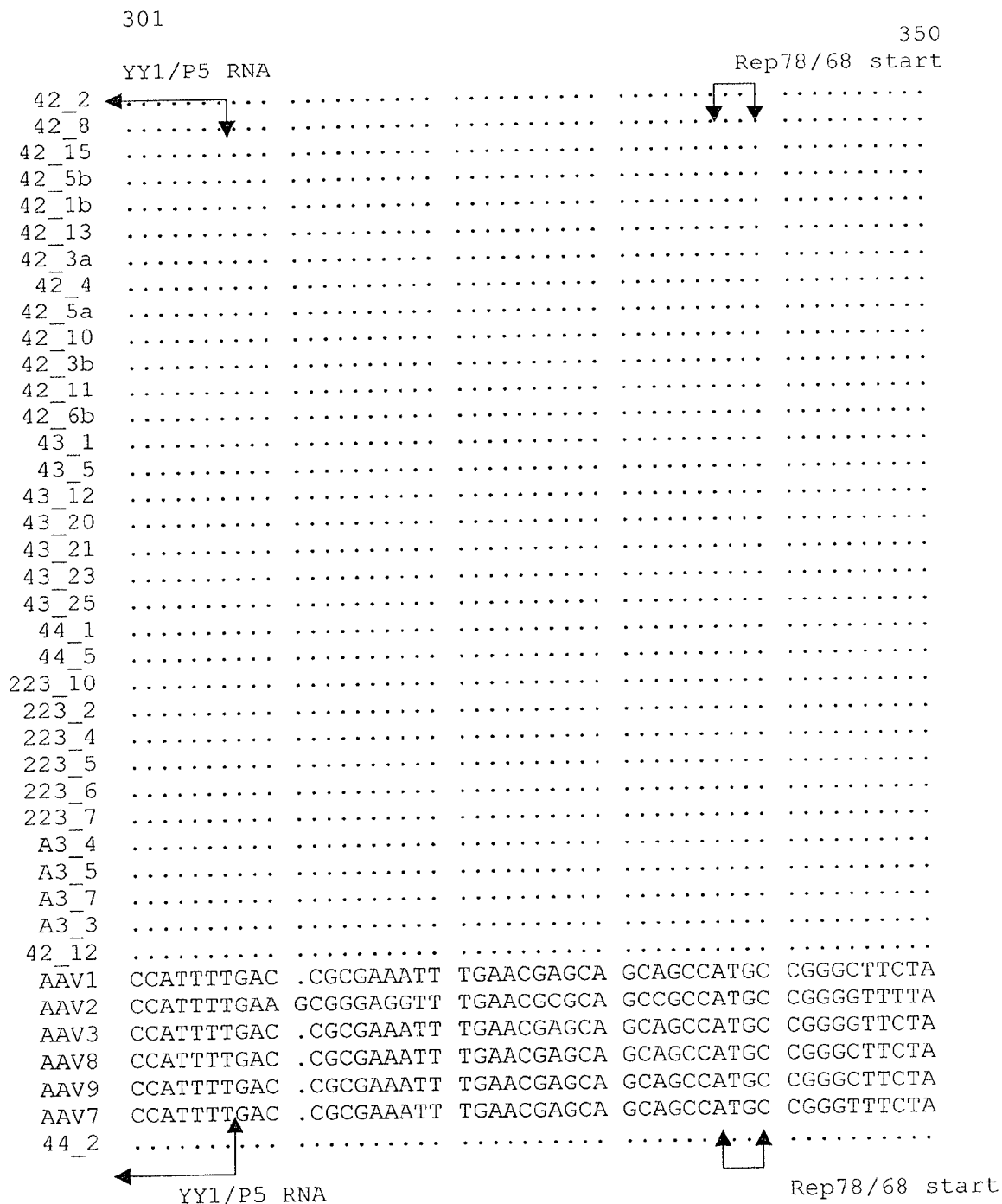


Fig. 1H

	351				400
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	CGAGATCGTG	ATCAAGGTGC	CGAGCGACCT	GGACGAGCAC	CTGCCGGGCA
AAV2	CGAGATTGTG	ATTAAGGTCC	CCAGCGACCT	TGACGGGCAT	CTGCCCGGCA
AAV3	CGAGATTGTC	CTGAAGGTCC	CGAGTGACCT	GGACGAGCGC	CTGCCGGGCA
AAV8	CGAGATCGTG	ATCAAGGTGC	CGAGCGACCT	GGACGAGCAC	CTGCCGGGCA
AAV9	CGAGATTGTG	ATCAAGGTGC	CGAGCGACCT	GGACGAGCAC	CTGCCGGGCA
AAV7	CGAGATCGTG	ATCAAGGTGC	CGAGCGACCT	GGACGAGCAC	CTGCCGGGCA
44_2	.....	.....	.....	.....	.....

Fig. 1I

	401				450
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	TTTCTGACTC	GTTTGTGAGC	TGGGTGGCCG	AGAAGGAATG	GGAGCTGCCC
AAV2	TTTCTGACAG	CTTTGTGAAC	TGGGTGGCCG	AGAAGGAATG	GGAGTTGCCC
AAV3	TTTCTAACTC	GTTTGTTAAC	TGGGTGGCCG	AGAAGGAATG	GGACGTGCCC
AAV8	TTTCTGACTC	GTTTGTGAAC	TGGGTGGCCG	AGAAGGAATG	GGAGCTGCCC
AAV9	TTTCTGACTC	TTTTGTGAAC	TGGGTGGCCG	AGAAGGAATG	GGAGCTGCCC
AAV7	TTTCTGACTC	GTTTGTGAAC	TGGGTGGCCG	AGAAGGAATG	GGAGCTGCCC
44_2	.....	.....	.....	.....	.....

Fig. 1J

	451				500
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	CCGGATTCTG	ACATGGATCT	GAATCTGATT	GAGCAGGCAC	CCCTGACCGT
AAV2	CCAGATTCTG	ACATGGATCT	GAATCTGATT	GAGCAGGCAC	CCCTGACCGT
AAV3	CCGGATTCTG	ACATGGATCC	GAATCTGATT	GAGCAGGCAC	CCCTGACCGT
AAV8	CCGGATTCTG	ACATGGATCG	GAATCTGATC	GAGCAGGCAC	CCCTGACCGT
AAV9	CCGGATTCTG	ACATGGATCG	GAATCTGATC	GAGCAGGCAC	CCCTGACCGT
AAV7	CCGGATTCTG	ACATGGATCT	GAATCTGATC	GAGCAGGCAC	CCCTGACCGT
44_2	.....	.....	.....	.....	.....

Fig. 1K

	501				550
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	GGCCGAGAAG	CTGCAGCGCG	ACTTCCTGGT	CCAATGGCGC	CGCGTGAGTA
AAV2	GGCCGAGAAG	CTGCAGCGCG	ACTTTCTGAC	GGAATGGCGC	CGTGTGAGTA
AAV3	GGCCGAAAAG	CTTCAGCGCG	AGTTCCTGGT	GGAGTGGCGC	CGCGTGAGTA
AAV8	GGCCGAGAAG	CTGCAGCGCG	ACTTCCTGGT	CCAATGGCGC	CGCGTGAGTA
AAV9	GGCCGAGAAG	CTGTAGCGCG	ACTTCCTGGT	CCAATGGCGC	CGCGTGAGTA
AAV7	GGCCGAGAAG	CTGCAGCGCG	ACTTCCTGGT	CCAATGGCGC	CGCGTGAGTA
44_2	.....	.....	.....	.....	.....

Fig. 1L

	551				600
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	AGGCCCCGGA	GGCCCTCTTC	TTTGTTTCAGT	TCGAGAAGGG	CGAGTCCTAC
AAV2	AGGCCCCGGA	GGCCCTTTTC	TTTGTGCAAT	TTGAGAAGGG	AGAGAGCTAC
AAV3	AGGCCCCGGA	GGCCCTCTTT	TTTGTCCAGT	TCGAAAAGGG	GGAGACCTAC
AAV8	AGGCCCCGGA	GGCCCTCTTC	TTTGTTTCAGT	TCGAGAAGGG	CGAGAGCTAC
AAV9	AGGCCCCGGA	GGCCCTCTTC	TTTGTTTCAGT	TCGAGAAGGG	CGAGAGCTAC
AAV7	AGGCCCCGGA	GGCCCTGTTC	TTTGTTTCAGT	TCGAGAAGGG	CGAGAGCTAC
44_2	.....	.....	.....	.....	.....

Fig. 1M

	601				650
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	TTCCACCTCC	ATATTCTGGT	GGAGACCACG	GGGGTCAAAT	CCATGGTGCT
AAV2	TTCCACATGC	ACGTGCTCGT	GGAAACCACC	GGGGTGAAAT	CCATGGTTTTT
AAV3	TTCCACCTGC	ACGTGCTGAT	TGAGACCATC	GGGGTCAAAT	CCATGGTGCT
AAV8	TTTCACCTGC	ACGTTCTGGT	CGAGACCACG	GGGGTCAAGT	CCATGGTGCT
AAV9	TTTCACCTGC	ACGTTCTGGT	CGAGACCACG	GGGGTCAAGT	CCATGGTGCT
AAV7	TTCCACCTTC	ACGTTCTGGT	GGAGACCACG	GGGGTCAAGT	CCATGGTGCT
44_2	.....	.....	.....	.....	.....

Fig. 1N

	651				700
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	GGGCCGCTTC	CTGAGTCAGA	TTAGGGACAA	GCT.GGTGCA	GACCATCTAC
AAV2	GGGACGTTTC	CTGAGTCAGA	TTCGCGAAAA	ACT..GATTC	AGAGAATTTA
AAV3	CGGCCGCTAC	GTGAGCCAGA	TTAAAGAGAA	GCT..GGTGA	CCCGCATCTA
AAV8	AGGCCGCTTC	CTGAGTCAGA	TTCGGGAAAA	GCTTGGTCCA	GACCATCTAC
AAV9	AGGCCGCTTC	CTGAGTCAGA	TTCGGGAGAA	GCT.GGTCCA	GACCATCTAC
AAV7	AGGCCGCTTC	CTGAGTCAGA	TTCGGGAGAA	GCT....G..	GTCCAGACCA
44_2	.....	.....	.....	.....	.....

Fig. 10

	701				750
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	C.GCGGGATC	GAGCCG.ACC	CTGCCCAACT	GGTTCGCGGT	GACCAA.GAC
AAV2	CCGCGGGATC	GAGCCG.ACT	TTGCCAAACT	GGTTCGCGGT	CACAAA...G
AAV3	CCGCGGGGTC	GAGCCG.CAG	CTTCCGAACT	GGTTCGCGGT	GACCAA...A
AAV8	CCGCGGGGTC	GAGCCCCACC	TTGCCCAACT	GGTTCGCGGT	GACCAAAGAC
AAV9	C.GCGGGATC	GAGCCG.ACC	CTGCCCAACT	GGTTCGCGGT	GACCAA.GAC
AAV7	TCTACCGCGG	GGTCGAGCCC	ACGCTGCCCA	ACTGGTTCGC	GGTGACCAAG
44_2	.....	.....	.....	.....	.....

Fig. 1P

	751				800
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	GCG.TAATGG	CGCCGGAGGG	GGG.AACAAG	GTGGTGGACG	AGTGCTACAT
AAV2	ACCAGAAATG	GCGCCGGAGG	CGGGAACAAG	GTGGTGGATG	AGTGCTACAT
AAV3	ACGCGAAATG	GCGCCGGGGG	CGGGAACAAG	GTGGTGGACG	ACTGCTACAT
AAV8	GCGGTAATGG	CGCCGGCGGG	GGGGAACAAG	GTGGTGGACG	AGTGCTACAT
AAV9	GCG.TAATGG	CGCCGGCGGG	GGG.AACAAG	GTGGTGGACG	AGTGCTACAT
AAV7	ACGCGTAATG	GCGCCGGCGG	GGGGAACAAG	GTGGTGGACG	AGTGCTACAT
44_2	.....	.....	.....	.....	.....

Fig. 1Q

	801				850
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	CCCCAACTAC	CTCCTGCCCCA	AGACTCAGCC	CGAGCTGCAG	TGGGCGTGGA
AAV2	CCCCAATTAC	TTGCTCCCCA	AAACCCAGCC	TGAGCTCCAG	TGGGCGTGGA
AAV3	CCCCAACTAC	CTGCTCCCCA	AGACCCAGCC	CGAGCTCCAG	TGGGCGTGGA
AAV8	CCCCAACTAC	CTCCTGCCCCA	AGACTCAGCC	CGAGCTGCAG	TGGGCGTGGA
AAV9	CCCCAACTAC	CTCCTGCCCCA	AGACTCAGCC	CGAGCTGCAG	TGGGCGTGGA
AAV7	CCCCAACTAC	CTCCTGCCCCA	AGACCCAGCC	CGAGCTGCAG	TGGGCGTGGA
44_2	.....	.....	.....	.....	.....

Fig. 1R

851

900

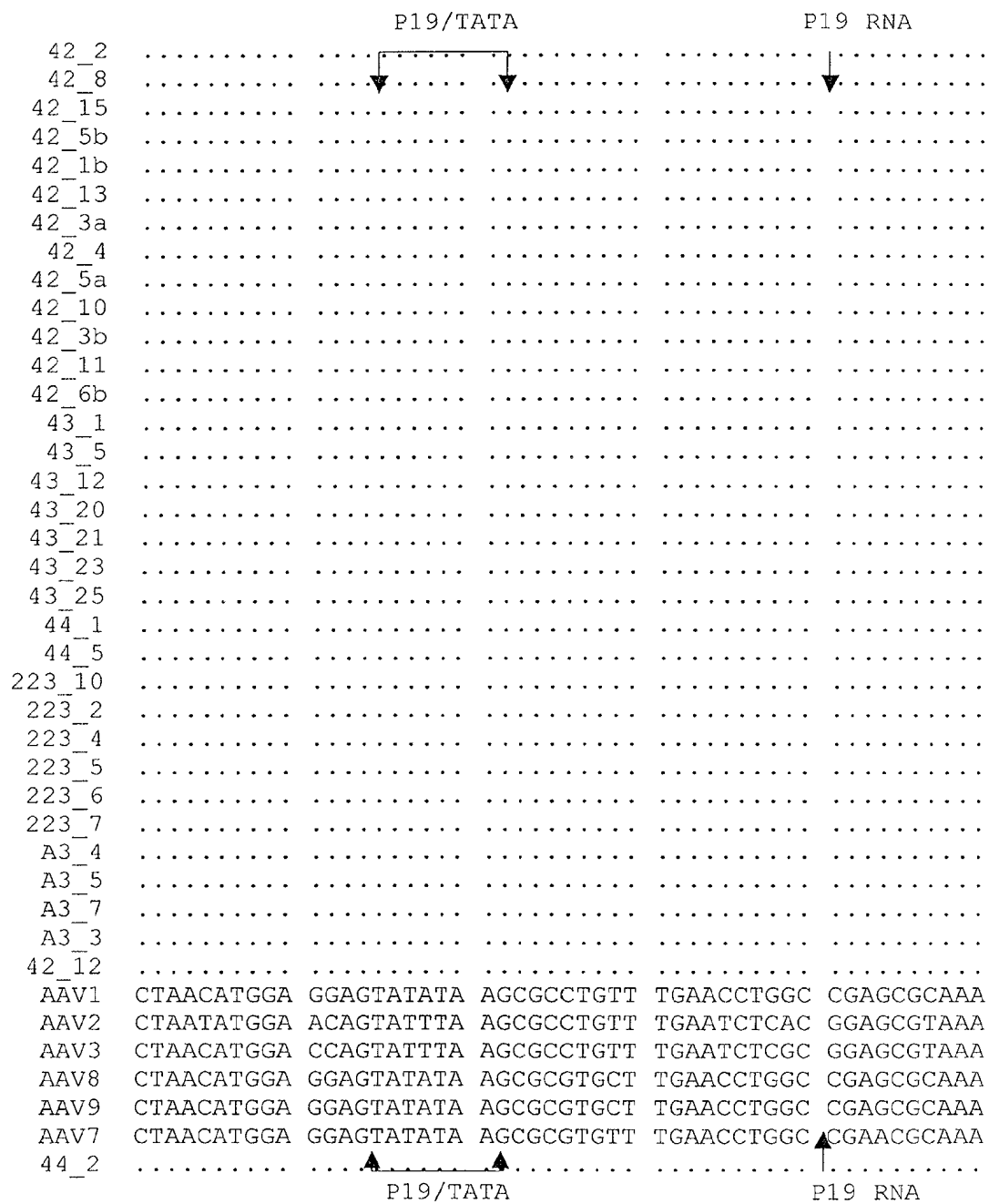


Fig. 1S

	901				950
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	CGGCTCGTGG	CGCAGCACCT	GACCCACGTC	AGCCAGACCC	AGGAGCAGAA
AAV2	CGGTTGGTGG	CGCAGCATCT	GACGCACGTG	TCGCAGACGC	AGGAGCAGAA
AAV3	CGGCTGGTGG	CGCAGCATCT	GACGCACGTG	TCGCAGACGC	AGGAGCAGAA
AAV8	CGGCTCGTGG	CGCAGCACCT	GACCCACGTC	AGCCAGACGC	AGGAGCAGAA
AAV9	CGGCTCGTGG	CGCAGCACCT	GACCCACGTC	AGCCAGACGC	AGGAGCAGAA
AAV7	CGGCTCGTGG	CGCAGCACCT	GACCCACGTC	AGCCAGACGC	AGGAGCAGAA
44_2	.....	.....	.....	.....	.....

Fig. 1T

	951				1000
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	CAAGGAGAAT	CTGAACCCCA	ATTCTGACGC	GCCTGTCATC	CGGTCAAAAA
AAV2	CAAAGAGAAT	CAGAATCCCA	ATTCTGATGC	GCCGGTGATC	AGATCAAAAA
AAV3	CAAAGAGAAT	CAGAACCCCA	ATTCTGACGC	GCCGGTCATC	AGGTCAAAAA
AAV8	CAAGGAGAAT	CTGAACCCCA	ATTCTGACGC	GCCCGTGATC	AGGTCAAAAA
AAV9	CAAGGAGAAT	CTGAACCCCA	ATTCTGACGC	GCCCGTGATC	AGGTCAAAAA
AAV7	CAAGGAGAAT	CTGAACCCCA	ATTCTGACGC	GCCCGTGATC	AGGTCAAAAA
44_2	.....	.....	.....	.....	.....

Fig. 1U

	1001					1050
		Rep52/40 start codon				
42_2	.....	.....	↓ ↓	.....	.....	.....
42_8	.....	.....		.....	.....	.....
42_15	.....	.....		.....	.....	.....
42_5b	.....	.....		.....	.....	.....
42_1b	.....	.....		.....	.....	.....
42_13	.....	.....		.....	.....	.....
42_3a	.....	.....		.....	.....	.....
42_4	.....	.....		.....	.....	.....
42_5a	.....	.....		.....	.....	.....
42_10	.....	.....		.....	.....	.....
42_3b	.....	.....		.....	.....	.....
42_11	.....	.....		.....	.....	.....
42_6b	.....	.....		.....	.....	.....
43_1	.....	.....		.....	.....	.....
43_5	.....	.....		.....	.....	.....
43_12	.....	.....		.....	.....	.....
43_20	.....	.....		.....	.....	.....
43_21	.....	.....		.....	.....	.....
43_23	.....	.....		.....	.....	.....
43_25	.....	.....		.....	.....	.....
44_1	.....	.....		.....	.....	.....
44_5	.....	.....		.....	.....	.....
223_10	.....	.....		.....	.....	.....
223_2	.....	.....		.....	.....	.....
223_4	.....	.....		.....	.....	.....
223_5	.....	.....		.....	.....	.....
223_6	.....	.....		.....	.....	.....
223_7	.....	.....		.....	.....	.....
A3_4	.....	.....		.....	.....	.....
A3_5	.....	.....		.....	.....	.....
A3_7	.....	.....		.....	.....	.....
A3_3	.....	.....		.....	.....	.....
42_12	.....	.....		.....	.....	.....
AAV1	CCTCCGCGCG	CTACATGGAG	CTGGTCGGGT	GGCTGGTGGA	CCGGGGGCATC	
AAV2	CTTCAGCCAG	GTACATGGAG	CTGGTCGGGT	GGCTCGTGGA	CAAGGGGATT	
AAV3	CCTCAGCCAG	GTACATGGAG	CTGGTCGGGT	GGCTGGTGGA	CCGCGGGATC	
AAV8	CCTCCGCGCG	CTATATGGAG	CTGGTCGGGT	GGCTGGTGGA	CCGGGGGCATC	
AAV9	CCTCCGCGCG	CTACATGGAG	CTGGTCGGGT	GGCTGGTGGA	CCGGGGGCATC	
AAV7	CCTCCGCGCG	CTACATGGAG	CTGGTCGGGT	GGCTGGTGGA	CCGGGGGCATC	
44_2	.....	.....	↑ ↑	.....	.....	.....
		Rep 52/40 start				

Fig. 1V

	1051				1100
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	ACCTCCGAGA	AGCAGTGGAT	CCAGGAGGAC	CAGGCCTCGT	ACATCTCCTT
AAV2	ACCTCCGAGA	AGCAGTGGAT	CCAGGAGGAC	CAGGCCTCAT	ACATCTCCTT
AAV3	ACGTCAGAAA	AGCAATGGAT	TCAGGAGGAC	CAGGCCTCGT	ACATCTCCTT
AAV8	ACCTCCGAGA	AGCAGTGGAT	CCAGGAGGAC	CAGGCCTCGT	ACATCTCCTT
AAV9	ACCTCCGAGA	AGCAGTGGAT	CCAGGAGGAC	CAGGCCTCGT	ACATCTCCTT
AAV7	ACCTCCGAGA	AGCAGTGGAT	CCAGGAGGAC	CAGGCCTCGT	ACATCTCCTT
44_2	.....	.....	.....	.....	.....

Fig. 1W

	1101				1150
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	CAACGCCGCT	TCCAACTCGC	GGTCCCAGAT	CAAGGCCGCT	CTGGACAATG
AAV2	CAATGCGGCC	TCCAACTCGC	GGTCCCAAAT	CAAGGCTGCC	TTGGACAATG
AAV3	CAACGCCGCC	TCCAACTCGC	GGTCCCAGAT	CAAGGCCGCG	CTGGACAATG
AAV8	CAACGCCGCC	TCCAACTCGC	GGTCCCAGAT	CAAGGCCGCG	CTGGACAATG
AAV9	CAACGCCGCC	TCCAACTCGC	GGTCCCAGAT	CAAGGCCGCG	CTGGACAATG
AAV7	CAACGCCGCC	TCCAACTCGC	GGTCCCAGAT	CAAGGCCGCG	CTGGACAATG
44_2	.....	.....	.....	.....	.....

Fig. 1X

	1151				1200
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	CCGGCAAGAT	CATGGCGCTG	ACCAAATCCG	CGCCCGACTA	CCTGGTAGGC
AAV2	CGGGAAAGAT	TATGAGCCTG	ACTAAAACCG	CCCCCGACTA	CCTGGTGGGC
AAV3	CCTCCAAGAT	CATGAGCCTG	ACAAAGACGG	CTCCGGACTA	CCTGGTGGGC
AAV8	CCGGCAAGAT	CATGGCGCTG	ACCAAATCCG	CGCCCGACTA	CCTGGTGGGG
AAV9	CCGGCAAGAT	CATGGCGCTG	ACCAAATCCG	CGCCCGACTA	CCTGGTAGGC
AAV7	CCGGCAAGAT	CATGGCGCTG	ACCAAATCCG	CGCCCGACTA	CCTGGTGGGG
44_2	.....	.....	.....	.....	.....

Fig. 1Y

	1201				1250
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....GA	ATTGCGCCCTT	TCTACGGCTG
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	CCCGCTCCGC	CCGCGGACAT	TAAAACCAAC	CGCATCTACC	GCATCCTGGA
AAV2	CAGCAGCCCG	TGGAGGACAT	TTCCAGCAAT	CGGATTTATA	AAATTTTGGA
AAV3	AGCAACCCGC	CGGAGGACAT	TACCAAAAAT	CGGATCTACC	AAATCCTGGA
AAV8	CCCTCGCTGC	CCGCGGACAT	TACCCAGAAC	CGCATCTACC	GCATCCTCGC
AAV9	CCTTCACTTC	CGGTGGACAT	TACGCAGAAC	CGCATCTACC	GCATCCTGCA
AAV7	CCCTCGCTGC	CCGCGGACAT	TAAAACCAAC	CGCATCTACC	GCATCCTGGA
44_2	.....	.....	.....	.....	.....

Fig. 1Z

	1251				1300
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	CGTCAACTGG	ACCAATGAGA	ACTTTCCTT	CAACGATTGC	GTCGACAAGA
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	GCTGAACGGC	TACGAACCTG	CCTACGCCGG	CTCCGTCTTT	CTCGGCTGGG
AAV2	ACTAAACGGG	TACGATCCCC	AATATGCGGC	TTCCGTCTTT	CTGGGATGGG
AAV3	GCTGAACGGG	TACGATCCGC	AGTACGCGGC	CTCCGTCTTC	CTGGGCTGGG
AAV8	TCTCAACGGC	TACGACCCTG	CCTACGCCGG	CTCCGTCTTT	CTCGGCTGGG
AAV9	GCTCAACGGC	TACGACCCTG	CCTACGCCGG	CTCCGTCTTT	CTCGGCTGGG
AAV7	GCTGAACGGG	TACGATCCTG	CCTACGCCGG	CTCCGTCTTT	CTCGGCTGGG
44_2	.....	.....	.....	.....	.....

Fig. 1AA

	1301				1350
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT	CGTGGAGTCC
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	CCCAGAAAAG	GTTCGGGAAG	CGCAACACCA	TCTGGCTGTT	TGGGCCGGCC
AAV2	CCACGAAAAA	GTTCGGCAAG	AGGAACACCA	TCTGGCTGTT	TGGGCCTGCA
AAV3	CGCAAAAGAA	GTTCGGGAAG	AGGAACACCA	TCTGGCTCTT	TGGGCCGGCC
AAV8	CTCAGAAAAA	GTTCGGGAAA	CGCAACACCA	TCTGGCTGTT	TGGACCCGCC
AAV9	CACAAAAGAA	GTTCGGGAAA	CGCAACACCA	TCTGGCTGTT	TGGGCCGGCC
AAV7	CCCAGAAAAA	GTTCGGGAAG	CGCAACACCA	TCTGGCTGTT	TGGGCCCGCC
44_2	.....	.....	.....	.....	.....

Fig. 1AB

	1351					1400
42_2	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
42_8	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
42_15	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
42_5b	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
42_1b	.....	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
42_3a	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
42_4	.....	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....GA	ATTGCCCTT
42_10	.....	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
42_6b	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC	AAAAGTGCAA	
43_1	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
43_5	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
43_12	.....	.....	.....	.....	.....GAA	TTCGCCCTT.
43_20	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
43_21	.....	.....	.....	.....	.....GAA	TTCGCCCTT.
43_23	.....	.....	.....	.....	.....GAA	TTCGCCCTT.
43_25	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
44_1	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
44_5	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
223_10	.....	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....GA	ATTGCCCTT
A3_5	.....	.....	.....	.....	.....GA	ATTGCCCTT
A3_7	.....	.....	.....A	GCGGCCGCGA	ATTGCCCTT	
A3_3	.....	.....	.....	.....	.....GA	ATTGCCCTT
42_12	.....	.....	.....	.....	.....GAA	TTCGCCCTTT
AAV1	ACCACGGGCA	AGACCAACAT	CGCGGAAGCC	ATCGCCACG	CCGTGCCCTT	
AAV2	ACTACGGGGA	AGACCAACAT	CGCGGAGGCC	ATAGCCACA	CTGTGCCCTT	
AAV3	ACGACGGGTA	AAACCAACAT	CGCGGAAGCC	ATCGCCACG	CCGTGCCCTT	
AAV8	ACCACGGGCA	AGACCAACAT	TGCGGAAGCC	ATCGCCACG	CCGTGCCCTT	
AAV9	ACCACGGGAA	AGACCAACAT	CGCAGAAGCC	ATTGCCACG	CCGTGCCCTT	
AAV7	ACCACGGGCA	AGACCAACAT	TGCGGAAGCC	ATCGCCACG	CCGTGCCCTT	
44_2	.....	.....	.....	.....	.....GA	ATTGCCCTT

Fig. 1AC

	1401				1450
42_2	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
42_8	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
42_15	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
42_5b	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
42_1b	.....	.....	.....	.....	.....
42_13	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
42_3a	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
42_4	.....	.....	.....	.....	.....
42_5a	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
42_6b	.GTCTTCCGC	CCAGATCGAT	CCCACCCCG	TGATCGTCAC	TTCCAACACC
43_1	.CTACGGCTG	CATCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
43_5	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
43_12	.....GGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
43_20	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
43_21	.....GGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
43_23	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
43_25	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
44_1	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
44_5	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	TCTACGGCTG	CGTCAACTGG	ACCAATGAAA	ACTTTCCCTT	CAACGATTGC
A3_5	TCTACGGCTG	CGTCAACTGG	ACCAATGAAA	ACTTTCCCTT	CAACGATTGC
A3_7	TCTACGGCTG	CGTCAACTGG	ACCAATGAAA	ACTTTCCCTT	CAACGATTGC
A3_3	TCTACGGCTG	CGTCAACTGG	ACCAATGAAA	ACTTTCCCTT	CAACGATTGC
42_12	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
AAV1	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAATGATTGC
AAV2	.CTACGGGTG	CGTAAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGACTGT
AAV3	.CTACGGCTG	CGTAAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
AAV8	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAATGATTGC
AAV9	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
AAV7	.CTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC
44_2	TCTACGGCTG	CGTCAACTGG	ACCAATGAGA	ACTTTCCCTT	CAACGATTGC

Fig. 1AD

	1451				1500
42_2	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
42_8	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
42_15	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
42_5b	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
42_1b	.....	.....	.....	.....	.....
42_13	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
42_3a	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
42_4	.....	.....	.....	.....	.....
42_5a	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
42_6b	AACATGTGCG	CCGTGATTGA	CGGGAACAGC	ACCACCTTCG	AGCACCAGCA
43_1	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
43_5	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
43_12	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
43_20	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
43_21	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
43_23	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
43_25	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
44_1	GTCGACAAGA	TGTTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
44_5	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGAAAGATGA	CCGCCAAGGT
A3_5	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGAAAGATGA	CCGCCAAGGT
A3_7	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGAAAGATGA	CCGCCAAGGT
A3_3	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGAAAGATGA	CCGCCAAGGT
42_12	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
AAV1	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
AAV2	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGGAAGATGA	CCGCCAAGGT
AAV3	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
AAV8	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
AAV9	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
AAV7	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT
44_2	GTCGACAAGA	TGGTGATCTG	GTGGGAGGAG	GGCAAGATGA	CGGCCAAGGT

Fig. 1AE

	1501				1550
42_2	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
42_8	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
42_15	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
42_5b	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
42_1b	.....	.....	.....	.....	.....
42_13	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
42_3a	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
42_4	.....	.....	.....	.....	.....
42_5a	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
42_6b	GCCGTTGCAG	GACCGGATGT	TCAAATTTGA	ACTCACCCGC	CGTCTGGAGC
43_1	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
43_5	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
43_12	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
43_20	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGTGTGGACC
43_21	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGTGTGGACC
43_23	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGTGTGGACC
43_25	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGTGTGGACC
44_1	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAAGTG	CGCGTGGACC
44_5	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAAGTG	CGCGTGGACC
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	CGTGGAATCT	GCCAAAGCCA	TTCTGGGTGG	AAGCAAGGTT	CGTGTGGACC
A3_5	CGTGGAATCT	GCCAAAGCCA	TTCTGGGTGG	AAGCAAGGTT	CGTGTGGACC
A3_7	CGTGGAATCT	GCCAAAGCCA	TTCTGGGTGG	AAGCAAGGTT	CGTGTGGACC
A3_3	CGTGGAATCT	GCCAAAGCCA	TTCTGGGTGG	AGGCAAGGTT	CGTGTGGACC
42_12	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
AAV1	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
AAV2	CGTGGAGTCG	GCCAAAGCCA	TTCTCGGAGG	AAGCAAGGTG	CGCGTGGACC
AAV3	CGTGGAGAGC	GCCAAGGCCA	TTCTGGGCGG	AAGCAAGGTG	CGCGTGGACC
AAV8	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
AAV9	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
AAV7	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAGGTG	CGCGTGGACC
44_2	CGTGGAGTCC	GCCAAGGCCA	TTCTCGGCGG	CAGCAAAGTG	CGCGTGGACC

Fig. 1AF

	1551				1600
42_2	AAAAGTGCAA	GTCTTCCGCC	CAGATCGATC	CCACCCCCGT	GATCGTCACT
42_8	AAAAGTGCAA	GTCTTCCGCC	CAGATCGATC	CCACCCCCGT	GATCGTCACT
42_15	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC
42_5b	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC
42_1b	.....	.....	.....	.....	.....
42_13	AAAAGTGCAA	GTCGTCCGCC	CAGATCGATC	CCACCCCCGT	GATCGTCACT
42_3a	AAAAGTGCAA	GTCGTCCGCC	CAGATCGATC	CCACCCCCGT	GATCGTCACT
42_4	.....	.....	.....	.....	.....
42_5a	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	AAAAGTGCAA	GTCTTCCGCC	CAGATCGATC	CCACCCCCGT	GATCGTCACT
42_6b	ATGACTTTGG	CAAGGTGACA	AAGCAGGAAG	TCAAAGAGTT	CTTCCGCTGG
43_1	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC
43_5	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC
43_12	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC
43_20	AAAAGTGCAA	GTCTTCCGCC	CAGATCGATC	CCACCCCCGT	GATCGTCACC
43_21	AAAAGTGCAA	GTCTTCCGCC	CAGATCGATC	CCACCCCCGT	GATCGTCACC
43_23	AAAAGTGCAA	GTCTTCCGCC	CAGATCGATC	CCACCCCCGT	GATCGTCACC
43_25	AAAAGTGCAA	GTCTTCCGCC	CAGATCGATC	CCACCCCCGT	GATCGTCACC
44_1	AAAAGTGCAA	GCCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC
44_5	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	AGAAATGCAA	GTCTTCCGCC	CAGATCGACC	CGACTCCGGT	GATTGTCACC
A3_5	AGAAATGCAA	GTCTTCCGCC	CAGATCGACC	CGACTCCGGT	GATTGTCACC
A3_7	AGAAATGCAG	GTCTTCCGCC	CAGATCGACC	CGACTCCGGT	GATTGTCACC
A3_3	AGAAATGCAA	GTCTTCCGCC	CAGATCGACC	CGACTCCGGT	GATTGTCACC
42_12	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC
AAV1	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC
AAV2	AGAAATGCAA	GTCTTCCGCC	CAGATAGACC	CGACTCCCGT	GATCGTCACC
AAV3	AAAAGTGCAA	GTCATCGGCC	CAGATCGAAC	CCACTCCCGT	GATCGTCACC
AAV8	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC
AAV9	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACTCCCGT	GATCGTCACC
AAV7	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC
44_2	AAAAGTGCAA	GTCGTCCGCC	CAGATCGACC	CCACCCCCGT	GATCGTCACC

Fig. 1AG

	1601				1650
42_2	TCCAACACCA	ACATGTGCGC	TGTGATTGAC	GGGAACAGCA	CCACCTTCGA
42_8	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
42_15	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
42_5b	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
42_1b	.....	.....	.....	.....	.....
42_13	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
42_3a	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
42_4	.....	.....	.....	.....	.....
42_5a	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
42_6b	GCGCAGGATC	ACGTGACCGA	GGTGGCGCAT	GAGTCTACG	TCAGAAAGGG
43_1	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
43_5	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
43_12	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
43_20	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCG	CCACCTTCGA
43_21	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
43_23	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
43_25	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
44_1	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
44_5	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	TCTAACACCA	ACATGTGCGC	CGTGATTGAC	GGAAACTCGA	CCACCTTCGA
A3_5	TCTAACACCA	ACATGTGCGC	CGTGATTGAC	GGAAACTCGA	CCACCTTCGA
A3_7	TCTAACACCA	ACATGTGCGC	CGTGATTGAC	GGAAACTCGA	CCACCTTCGA
A3_3	TCTAACACCA	ACATGTGCGC	CGTGATTGAC	GGAAACTCGA	CCACCTTCGA
42_12	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
AAV1	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
AAV2	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACCAA	CGACCTTCGA
AAV3	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
AAV8	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
AAV9	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
AAV7	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA
44_2	TCCAACACCA	ACATGTGCGC	CGTGATTGAC	GGGAACAGCA	CCACCTTCGA

Fig. 1AH

	1651				1700
42_2	GCACCAGCAG	CCGTTACAAG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
42_8	GCACCAGCAG	CCGTTACAAG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
42_15	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
42_5b	GCACCAGCAG	CCGTTACAAG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
42_1b	.....	.....	.....	.....	.....
42_13	GCACCAGCAG	CCGTTACAAG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
42_3a	GCACCAGCAG	CCGTTACAAG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
42_4	.....	.....	.....	.....	.....
42_5a	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	GCACCAGCAG	CCGTTACAAG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
42_6b	TGGAGCCAAC	AAGAGACCCG	CCCCGATGA	CGCGGATAAA	AGCGAGCCCA
43_1	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAGTTCGAA	CTCACCCGCC
43_5	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAGTTCGAA	CTCACCCGCC
43_12	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAGTTCGAA	CTCACCCGCC
43_20	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
43_21	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
43_23	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
43_25	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
44_1	GCACCAGCAG	CCGTTGCGGG	ACCGGATGTT	CAAGTTTGAA	CTCACCCGCC
44_5	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAGTTTGAA	CTCACCCGCC
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	GCACCAGCAG	CCGTTGCAAG	ACCGGATGTT	CAAATTTGAA	CTTACCCGCC
A3_5	GCACCAGCAG	CCGTTGCAAG	ACCGGATGTT	CAAATTTGAA	CTTACCCGCC
A3_7	GCACCAGCAG	CCGTTGCAAG	ACCGGATGTT	CAAATTTGAA	CTTACCCGCC
A3_3	GCACCAGCAG	CCGTTGCAAG	ACCGGATGTT	CAAATTTGAA	CTTACCCGCC
42_12	GCACCAGCAG	CCGTTACAAG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
AAV1	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
AAV2	ACACCAGCAG	CCGTTGCAAG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
AAV3	GCATCAGCAG	CCGCTGCAGG	ACCGGATGTT	TGAATTTGAA	CTTACCCGCC
AAV8	GCACCAGCAG	CCTCTCCAGG	ACCGGATGTT	TAAGTTCGAA	CTCACCCGCC
AAV9	GCACCAGCAG	CCTCTCCAGG	ACCGGATGTT	TAAGTTCGAA	CTCACCCGCC
AAV7	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAATTTGAA	CTCACCCGCC
44_2	GCACCAGCAG	CCGTTGCAGG	ACCGGATGTT	CAAGTTTGAA	CTCACCCGCC

Fig. 1AI

	1701					1750
42_2	GTCTGGAGCA	CGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAAAGAGTTC	
42_8	GTCTGGAGCA	CGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAAAGAGTTC	
42_15	GTCTGGAGCA	TGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAAAGAGTTC	
42_5b	GTCTGGAGCA	CGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAAAGAGTTC	
42_1b	.....	.....	.....	.....	.....	
42_13	GTCTGGAGCA	TGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAAAGAGTTC	
42_3a	GTCTGGAGCA	TGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAAAGAGTTC	
42_4	.....	.....	.....	.....	.....	
42_5a	GTCTGGAGCA	TGACTTTGGC	AAGGCGACAA	AGCAGGAAGT	CAAAGAGTTC	
42_10	.....	.....	.....	.....	.....	
42_3b	.....	.....	.....	.....	.....	
42_11	GTCTGGAGCA	CGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAAAGAGTTC	
42_6b	AGCGGGCCTG	CCCCTCAGTC	GCGGATCCAT	CGACGTCAGA	CGCGGAAGGA	
43_1	GTCTGGAGCA	CGACTTTGGC	AAGGTGACCA	AGCAGGAAGT	CAAAGAGTTC	
43_5	GTCTGGAGCA	CGACTTTGGC	AAGGTGACCA	AGCAGGAAGT	CAAAGAGTTC	
43_12	GTCTGGAGCA	CGACTTTGGC	AAGGTGACCA	AGCAGGAAGT	CAAAGAGTTC	
43_20	GTCTGGAGCA	TGACTTTGGC	AAGGTGACGA	AGCAGGAAGT	CAAAGAGTTC	
43_21	GTCTGGAGCA	TGACTTTGGC	AAGGTGACGA	AGCAGGAAGT	CAAAGAGTTC	
43_23	GTCTGGAGCA	TGACTTTGGC	AAGGTGACGA	AGCAGGAAGT	CAAAGAGTTC	
43_25	GTCTGGAGCA	TGACTTTGGC	AAGGTGACGA	AGCAGGAAGT	CAAAGGGTTC	
44_1	GTCTGGAGCA	CGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAGAGAGTTC	
44_5	GTCTGGAGCA	CGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAGAGAGTTC	
223_10	.....	.....	.....	.....	.....	
223_2	.....	.....	.....	.....	.....	
223_4	.....	.....	.....	.....	.....	
223_5	.....	.....	.....	.....	.....	
223_6	.....	.....	.....	.....	.....	
223_7	.....	.....	.....	.....	.....	
A3_4	GTTTGGATCA	TGACTTTGGG	AAGGTCACCA	AGCAGGAAGT	CAAAGACTTT	
A3_5	GTTTGGATCA	TGACTTTGGG	AAGGTCACCA	AGCAGGAAGT	CAAAGACTTT	
A3_7	GTTTGGATCA	TGACTTTGGG	AAGGTCACCA	AGCAGGAAGT	CAAAGACTTT	
A3_3	GTTTGGATCA	TGACTTTGGG	AAGGTCACCA	AGCAGGAAGT	CAAAGACTTT	
42_12	GTCTGGAGCA	CGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAAAGAGTTC	
AAV1	GTCTGGAGCA	TGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAAAGAGTTC	
AAV2	GTCTGGATCA	TGACTTTGGG	AAGGTCACCA	AGCAGGAAGT	CAAAGACTTT	
AAV3	GTTTGGACCA	TGACTTTGGG	AAGGTCACCA	AACAGGAAGT	AAAGGACTTT	
AAV8	GTCTGGAGCA	CGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAAAGAGTTC	
AAV9	GTCTGGAGCA	CGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAAAGAGTTC	
AAV7	GTCTGGAGCA	CGACTTTGGC	AAGGTGACGA	AGCAGGAAGT	CAAAGAGTTC	
44_2	GTCTGGAGCA	CGACTTTGGC	AAGGTGACAA	AGCAGGAAGT	CAGAGAGTTC	

Fig. 1AJ

	1751				1800
42_2	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
42_8	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
42_15	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
42_5b	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
42_1b	.....	.....	.....	.....	.....
42_13	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
42_3a	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
42_4	.....	.....	.....	.....	.....
42_5a	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
42_6b	GCTCCGGTGG	ACTTTGCCGA	CAGGTACCAA	AACAAATGTT	CTCGTCACGC
43_1	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
43_5	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
43_12	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
43_20	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCCACGT
43_21	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCCACGT
43_23	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCCACGT
43_25	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCCACGT
44_1	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCACG	AGTTCTACGT
44_5	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCACG	AGTTCTACGT
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	TTCCGGTGGG	CTCAAGATCA	CGTGACTGAG	GTGGAGCATG	AGTTCTACGT
A3_5	TTCCGGTGGG	CTCAAGATCA	CGTGACTGAG	GTGGAGCATG	AGTTCTACGT
A3_7	TTCCGGTGGG	CTCAAGATCA	CGTGACTGAG	GTGGAGCATG	AGTTCTACGT
A3_3	TTCCGGTGGG	CTCAAGATCA	CGTGACTGAG	GTGGAGCATG	AGTTCTACGT
42_12	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
AAV1	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
AAV2	TTCCGGTGGG	CAAAGGATCA	CGTGGTTGAG	GTGGAGCATG	AATTCTACGT
AAV3	TTCCGGTGGG	CTTCCGATCA	CGTGACTGAC	GTGGCTCATG	AGTTCTACGT
AAV8	TTCCGCTGGG	CCAGTGATCA	CGTGACCGAG	GTGGCGCATG	AGTTTTACGT
AAV9	TTCCGCTGGG	CCAGTGATCA	CGTGACCGAG	GTGGCGCATG	AGTTTTACGT
AAV7	TTCCGCTGGG	CCAGTGATCA	CGTGACCGAG	GTGGCGCATG	AGTTCTACGT
44_2	TTCCGCTGGG	CGCAGGATCA	CGTGACCGAG	GTGGCGCACG	AGTTCTACGT

Fig. 1AK

					1801		1850
							P40/TATA
42_2	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
42_8	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
42_15	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
42_5b	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
42_1b	.....	.....	.....	.....	.....	.....	
42_13	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
42_3a	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
42_4	.....	.....	.....	.....	.....	.....	
42_5a	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
42_10	.....	.....	.....	.....	.....	.....	
42_3b	.....	.....	.....	.....	.....	.....	
42_11	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
42_6b	GGGCATAGCG	CTGACGTAAA	TCACGTCATA	GGGGAGTGGT	CCTGTATTAG		
43_1	CAGAAAGGGC	GGAGCCAGCA	AAAGACCCGC	CCCCGATGAC	GCGGATATAA		
43_5	CAGAAAGGGC	GGAGCCAGCA	AAAGACCCGC	CCCCGATGAC	GCGGATATAA		
43_12	CAGAAAGGGC	GGAGCCAGCA	AAAGACCCGC	CCCCGATGAC	GCGGATATAA		
43_20	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGATATAA		
43_21	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGATATAA		
43_23	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGATATAA		
43_25	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGATATAA		
44_1	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
44_5	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
223_10	.....	.....	.....	.....	.....	.....	
223_2	.....	.....	.....	.....	.....	.....	
223_4	.....	.....	.....	.....	.....	.....	
223_5	.....	.....	.....	.....	.....	.....	
223_6	.....	.....	.....	.....	.....	.....	
223_7	.....	.....	.....	.....	.....	.....	
A3_4	CAAAAAGGGT	GGAGCCAAGA	AAAGGCCCGC	CCCCGATGAT	GTATATATAA		
A3_5	CAAAAAGGGT	GGAGCCAAGA	AAAGGCCCGC	CCCCGATGAT	GTATATATAA		
A3_7	CAAAAAGGGT	GGAGCCAAGA	AAAGGCCCGC	CCCCGATGAT	GTATATATAA		
A3_3	CAAAAAGGGT	GGAGCCAAGA	AAAGGCCCGC	CCCCGATGAT	GTATATATAA		
42_12	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
AAV1	CAGAAAGGGT	GGAGCCAACA	AAAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
AAV2	CAAAAAGGGT	GGAGCCAAGA	AAAGACCCGC	CCCCAGTGAC	GCAGATATAA		
AAV3	CAGAAAGGGT	GGAGCTAAGA	AACGCCCCGC	CTCCAATGAC	GCGGATGTAA		
AAV8	CAGAAAGGGC	GGAGCCAGCA	AAAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
AAV9	CAGAAAGGGC	GGAGCCAGCA	AAAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
AAV7	CAGAAAGGGC	GGAGCCAGCA	AAAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
44_2	CAGAAAGGGT	GGAGCCAACA	AGAGACCCGC	CCCCGATGAC	GCGGAT	AAAA	
							P40/TATA

Fig. 1AL

	1851				1900
				P40 RNA	
42_2	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
42_8	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
42_15	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
42_5b	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
42_1b	.....	.....	.....	.....	.....
42_13	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
42_3a	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
42_4	.....	.....	.....	.....	.....
42_5a	GCGAGCCCAA	GCGGGCCCGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
42_6b	CTGTACAGTG	AGTGCTTTTG	CGACATTTTG	C..ATCCATC	GACGTCAGAC
43_1	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
43_5	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
43_12	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
43_20	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
43_21	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
43_23	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
43_25	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
44_1	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
44_5	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	ATGAGCCCAA	GCGGGCGCGC	GAGTCAGTTG	CGCAGCCATC	GACGTCAGAC
A3_5	ATGAGCCCAA	GCGGGCGCGC	GAGTCAGTTG	CGCAGCCATC	GACGTCAGAC
A3_7	ATGAGCCCAA	GCGGGCGCGC	GAGTCAGTTG	CGCAGCCATC	GACGTCAGAC
A3_3	ATGAGCCCAA	GCGGGCGCGC	GAGTCAGTTG	CGCAGCCATC	GACGTCAGAC
42_12	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
AAV1	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
AAV2	GTGAGCCCAA	ACGGGTGCGC	GAGTCAGTTG	CGCAGCCATC	GACGTCAGAC
AAV3	GCGAGCCAAA	ACGGGAGTGC	ACGTCACTTG	CGCAGCCGAC	AACGTCAGAC
AAV8	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
AAV9	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
AAV7	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
44_2	GCGAGCCCAA	GCGGGCCTGC	CCCTCAGTCG	CGGATCCATC	GACGTCAGAC
				P40 RNA	

Fig. 1AM

	1901				1950
42_2	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
42_8	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
42_15	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
42_5b	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
42_1b	.....	.....	.....	.....	.....
42_13	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
42_3a	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
42_4	.....	.....	.....	.....	.....
42_5a	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
42_6b	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
43_1	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
43_5	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
43_12	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
43_20	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
43_21	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
43_23	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
43_25	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
44_1	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
44_5	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	GCGGA...AG	CTTCGATAAA	CTACGCGGGC	AGGTACCAAA	ACAAATGTTC
A3_5	GCGGA...AG	CTTCGATAAA	CTACGCGGAC	AGGTACCAAA	ACAAATGTTC
A3_7	GCGGA...AG	CTTCGATAAA	CTACGCGGAC	AGGTACCAAA	ACAAATGTTC
A3_3	GCGGA...AG	CTTCGATAAA	CTACGCGGAC	AGGTACCAAA	ACAAATGTTC
42_12	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
AAV1	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
AAV2	GCGGA...AG	CTTCGATCAA	CTACGCAGAC	AGGTACCAAA	ACAAATGTTC
AAV3	GCGGA...AG	CACCGGCGGA	CTACGCGGAC	AGGTACCAAA	ACAAATGTTC
AAV8	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
AAV9	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
AAV7	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC
44_2	GCGGAAGGAG	CTCCGGTGGA	CTTTGCCGAC	AGGTACCAAA	ACAAATGTTC

Fig. 1AN

	1951				2000
42_2	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
42_8	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
42_15	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
42_5b	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
42_1b	.....	.....	....GAATTC	GCCCTT....	.GGCTGCGTC
42_13	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
42_3a	TCGTCACGCG	GGCATGCTTC	AGATGCTGCT	TCCCTG.CAA	GACATGCGAG
42_4	.....	.....	....GAATTC	GCCCTTTCTA	CGGCTGCGTC
42_5a	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	AACATGCGAG
42_10	.....	.....	....GAATTC	GCCCTTTCTA	CGGCTGCGTC
42_3b	.....	.....	....GAATTC	GCCCTTTCTA	CGGCTGCGTC
42_11	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
42_6b	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
43_1	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	AACGTGCGAG
43_5	TCGTCACGCG	GGCATGCTTC	AGACGCTGTT	TCCCTG.CAA	AACGTGCGAG
43_12	TCGTCACGCG	GGCATGCTCC	AGATGCTGTT	TCCCTG.CAA	AACGTGCGAG
43_20	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
43_21	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
43_23	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
43_25	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
44_1	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	AACATGCGAG
44_5	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	AACATGCGAG
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	TCGTCACGTG	GGCATGAATC	TGATGCTGTT	TCCCTG.TCG	ACAATGCGAA
A3_5	TCGTCACGTG	GGCATGAATC	TGATGCTGTT	TCCCTG.TCG	ACAATGCGAA
A3_7	TCGTCACGTG	GGCATGAATC	TGATGCTGTT	TCCCTG.TCG	ACAATGCGAA
A3_3	TCGTCACGTG	GGCATGAATC	TGATGCTGTT	TCCCTG.TCG	ACAATGCGAA
42_12	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
AAV1	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	GACATGCGAG
AAV2	TCGTCACGTG	GGCATGAATC	TGATGCTGTT	TCCCTG.CAG	ACAATGCGAG
AAV3	TCGTCACGTG	GGCATGAATC	TGATGCTTTT	TCCCTG.TAA	AACATGCGAG
AAV8	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	AACGTGCGAG
AAV9	TCGTCACGCG	GGCATGCTTC	AGATGCTGCT	TCCCTG.CAA	AACGTGCGAG
AAV7	TCGTCACGCG	GGCATGATTC	AGATGCTGTT	TCCCTG.CAA	AACGTGCGAG
44_2	TCGTCACGCG	GGCATGCTTC	AGATGCTGTT	TCCCTG.CAA	AACATGCGAG

Fig. 1AO

	2001				2050
42_2	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
42_8	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
42_15	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCGCGGGA	CCAGAGACTG
42_5b	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
42_1b	A.ACTGGACC	A..ATGAGAA	CTTTCCCTTC	A.....A	CGATTGCGTC
42_13	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
42_3a	AGAATGAATC	AGAATTTCAG	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
42_4	A.ACTGGACC	A..ATGAGAA	CTTTCCCTTC	A.....A	CGATTGCGTC
42_5a	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
42_10	A.ACTGGACC	A..ATGAGAA	CTTTCCCTTC	A.....A	CGATTGCGTC
42_3b	A.ACTAGACC	A..ATGAGAA	CTTTCCCTTC	A.....A	CGATTGCGTC
42_11	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCGAGAGACTG
42_6b	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
43_1	AAAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGG	TCAGAGACTG
43_5	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGG	TCAGAGACTG
43_12	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGG	TCAGAGACTG
43_20	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
43_21	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
43_23	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
43_25	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
44_1	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
44_5	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	AGAATGAATC	AGAATTCAAA	TATCTGCTTC	ACACACGGGC	AAAAAGACTG
A3_5	AGAATGAATC	AGAATTCAAA	TATCTGCTTC	ACACACGGGC	AAAAAGACTG
A3_7	AGAATGAATC	AGAATTCAAA	TATCTGCTTC	ACACACGGGC	AAAAAGACTG
A3_3	AGAATGAATC	AGAATTCAAA	TATCTGCTTC	ACACACGGGC	AAAAAGACTG
42_12	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG
AAV1	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CGAGAGACTG
AAV2	AGAATGAATC	AGAATTCAAA	TATCTGCTTC	ACTCACGGAC	AGAAAGACTG
AAV3	AGAATGAATC	AAATTTCCAA	TGTCTGTTTT	ACGCATGGTC	AAAGAGACTG
AAV8	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACACACGGGG	TCAGAGACTG
AAV9	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACACACGGGG	TCAGAGACTG
AAV7	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACACACGGGG	TCAGAGACTG
44_2	AGAATGAATC	AGAATTTCAA	CATTTGCTTC	ACGCACGGGA	CCAGAGACTG

Fig. 1AP

	2051				2100
42_2	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
42_8	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
42_15	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
42_5b	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
42_1b	GACAAGATGG	TGATCTGGTG	GG..AGGAGG	GCAAGA....	..TGACGGCC
42_13	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
42_3a	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
42_4	GACAAGATGG	TGATCTGGTG	GG..AGGAGG	GCAAGA....	..TGACGGCC
42_5a	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
42_10	GACAAGATGG	TGATCTGGTG	GG..AGGAGG	GCAAGA....	..TGACGGCC
42_3b	GACAAGATGG	TGATCTGGTG	GG..AGGAGG	GCAAGA....	..TGACGGCC
42_11	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
42_6b	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
43_1	CTCAGAATGT	TTCCCCGGTG	CATCAGAATC	TCAACC....	..GGTCGTCA
43_5	CTCAGAATGT	TTCCCCGGTG	CATCAGAATC	TCAACC....	..GGTCGTCA
43_12	CTCAGAATGT	TTCCCCGGTG	CATCAGAATC	TCAACC....	..GGTCGTCA
43_20	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
43_21	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
43_23	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
43_25	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
44_1	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
44_5	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTTGTCA
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	TTTGGGAATGC	TTTCCCG...	TGTCAGAATC	TCAACCCGTT	TCTGTCTGTCA
A3_5	TTTGGGAATGC	TTTCCCG...	TGTCAGAATC	TCAACCCGTT	CCTGTCTGTCA
A3_7	TTTGGGAATGC	TTTCCCG...	TGTCAGAATC	TCAACCCGTT	TCTGTCTGTCA
A3_3	TTTGGGAATGC	TTTCCCG...	TGTCAGAATC	TCAACCCGTT	TCTGTCTGTCA
42_12	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
AAV1	TTCAGAGTGC	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
AAV2	TTTAGAGTGC	TTTCCCG...	TGTCAGAATC	TCAACCCGTT	TCTGTCTGTCA
AAV3	TGGGGAATGC	TTCCCTGGAA	TGTCAGAATC	TCAACCCGTT	TCTGTCTGTCA
AAV8	CTCAGAGTGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
AAV9	CTCAGAGTGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
AAV7	TTTAGAGTGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA
44_2	TTCAGAATGT	TTCCCCGGCG	TGTCAGAATC	TCAACC....	..GGTCGTCA

Fig. 1AQ

	2101				2150
42_2	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
42_8	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTAGGG.CG
42_15	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
42_5b	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
42_1b	.AAGGTCGTG	GAGTCCGCCA	AG...GCCA	TTCATCATCT	GCTGGGG.CG
42_13	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
42_3a	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
42_4	.AAGGTCGTG	GAGTCCGCCA	AG...GCCA	TTCATCATCT	GCTGGGG.CG
42_5a	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
42_10	AA...GGTC	GTGAAGTCCG	CCAAG.GCCA	TTCATCATCT	GCTGGGG.CG
42_3b	AA...GGTC	GTGGAGTCCG	CCAAG.GCCA	TTCATCATCT	GCTGGGG.CG
42_11	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
42_6b	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
43_1	GAAAAAAAAC	GTATCAGAAA	CTGTGTGCCA	TTCATCATCT	GCTGGGG.CG
43_5	GAAAAAAAAC	GTATCAGAAA	CTGTGTGCCA	TTCATCATCT	GCTGGGG.CG
43_12	GAAAAAAAAC	GTATCAGAAA	CTGTGTGCCA	TTCATCATCT	GCTGGGG.CG
43_20	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
43_21	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
43_23	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
43_25	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
44_1	GAAAAAAGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
44_5	GAAAAAAGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	GAAAAACG..	.TATCAGAAA	CTTTGTTACA	TTCATCATAT	CATGGGA.AA
A3_5	GAAAAACG..	.TATCAGAAA	CTTTGTTACA	TTCATCATAT	CATGGGA.AA
A3_7	GAAAAACG..	.TATCAGAAA	CTTTGTTACA	TTCATCATAT	CATGGGA.AA
A3_3	GAAAAACG..	.TATCAGAAA	CTTTGTTACA	TTCATCATAT	CATGGGA.AA
42_12	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
AAV1	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
AAV2	AAAAGGCG..	.TATCAGAAA	CTGTGCTACA	TTCATCATAT	CATGGGA.AA
AAV3	AAAAGAAGAC	TTATCAGAAA	CTGTGTCCAA	TTCATCATAT	CCTGGGA.AG
AAV8	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
AAV9	GAAAGAGGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGG.CG
AAV7	GAAAAAAGAC	GTATCGGAAA	CTCTGCGCGA	TTCATCATCT	GCTGGGG.CG
44_2	GAAAAAAGAC	GTATCGGAAA	CTCTGTGCCA	TTCATCATCT	GCTGGGGGCG

Fig. 1AR

	2151				2200
42_2	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
42_8	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
42_15	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
42_5b	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
42_1b	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
42_13	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
42_3a	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
42_4	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
42_5a	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
42_10	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
42_3b	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
42_11	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
42_6b	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
43_1	GGCACCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
43_5	GGCACCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
43_12	GGCACCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
43_20	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
43_21	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
43_23	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
43_25	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
44_1	GGCACCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTAG
44_5	GGCACCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTAG
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	AGAACCAGAC	...GCCTGCA	CTGCCTGCGA	CCTGGTAAAT	GTGGACTTGG
A3_5	AGTACCAGAC	...GCCTGCA	CTGCCTGCGA	CCTGGTAAAT	GTGGACTTGG
A3_7	AGTACCAGAC	...GCCTGCA	CTGCCTGCGA	CCTGGTAAAT	GTGGACTTGG
A3_3	AGTACCAGAC	...GCCTGCA	CTGCCTGCGA	CCTGGTAAAT	GTGGACTTGG
42_12	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
AAV1	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
AAV2	GGTGCCAGAC	...GCTTGCA	CTGCCTGCGA	TCTGGTCAAT	GTGGATTGG
AAV3	GGCACCCGAG	ATTGCCTGTT	CGGCCTGCGA	TTTGGCCAAT	GTGGACTTGG
AAV8	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
AAV9	GGCTCCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTGG
AAV7	GGCGCCCGAG	ATTGCTTGCT	CGGCCTGCGA	CCTGGTCAAC	GTGGACCTGG
44_2	GGCACCCGAG	ATTGCTTGCT	CGGCCTGCGA	TCTGGTCAAC	GTGGACCTAG

Fig. 1AS

2201

2250

			<u>Rep 78 stop</u>	<u>vp1 start</u>
42_2	ATGACCGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
42_8	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
42_15	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
42_5b	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
42_1b	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
42_13	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
42_3a	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
42_4	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
42_5a	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
42_10	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
42_3b	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
42_11	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
42_6b	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
43_1	ACGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
43_5	ACGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
43_12	ACGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
43_20	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
43_21	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
43_23	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
43_25	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
44_1	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
44_5	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
223_10	.....	.....	.....	.....
223_2	.....	.....	.....	.....
223_4	.....	.....	.....	.....
223_5	.....	.....	.....	.....
223_6	.....	.....	.....	.....
223_7	.....	.....	.....	.....
A3_4	ATGACTGTAT	TTCTGAGCAA	TAAATGACTT	AAATCAGGTA TGGCTGCTGA
A3_5	ATGACTGTAT	TTCTGAGCAA	TAAATGACTT	AAATCAGGTA TGGCTGCTGA
A3_7	ATGACTGTAT	TTCTGAGCAA	TAAATGACTT	AAATCAGGTA TGGCTGCTGA
A3_3	ATGACTGTAT	TTCTGAGCAA	TAAATGACTT	AAATCAGGTA TGGCTGCTGA
42_12	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
AAV1	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
AAV2	ATGACTGCAT	CTTTGAACAA	TAAATGATTT	AAATCAGGTA TGGCTGCCGA
AAV3	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCTGA
AAV8	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
AAV9	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
AAV7	ACGACTGCGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA
44_2	ATGACTGTGT	TTCTGAGCAA	TAAATGACTT	AAACCAGGTA TGGCTGCCGA

Rep78 stop                      vp1 start

Fig. 1AT

2251						2300
						<u>Rep68 stop</u>
42_2	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
42_8	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
42_15	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
42_5b	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
42_1b	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
42_13	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
42_3a	TGGTCATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
42_4	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
42_5a	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
42_10	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
42_3b	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
42_11	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
42_6b	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
43_1	TGGTTATCTT	CCAGATTGGC	TTGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
43_5	TGGTTATCTT	CCAGATTGGC	TTGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
43_12	TGGTTATCTT	CCAGATTGGC	TTGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
43_20	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
43_21	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
43_23	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
43_25	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
44_1	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
44_5	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
223_10	.....	.....	.....	.....	.....	
223_2	.....	.....	.....	.....	.....	
223_4	.....	.....	.....	.....	.....	
223_5	.....	.....	.....	.....	.....	
223_6	.....	.....	.....	.....	.....	
223_7	.....	.....	.....	.....	.....	
A3_4	CGGTTATCTT	CCAGATTGGC	TCGAGGACAC	TCTCTCTGAA	GGAAATCAGAC	
A3_5	CGGTTATCTT	CCAGATTGGC	TCGAGGACAC	TCTCTCTGAA	GGAAATCAGAC	
A3_7	CGGTTATCTT	CCAGATTGGC	TCGAGGACAC	TCTCTCTGAA	GGAAATCAGAC	
A3_3	CGGTTATCTT	CCAGATTGGC	TCGAGGACAC	TCTCTCTGAA	GGAAATCAGAC	
42_12	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATCCGCG	
AAV1	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
AAV2	TGGTTATCTT	CCAGATTGGC	TCGAGGACAC	TCTCTCTGAA	GGAAATAAGAC	
AAV3	CGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTTTCTGAA	GGCATTCGTG	
AAV8	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
AAV9	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
AAV7	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
44_2	TGGTTATCTT	CCAGATTGGC	TCGAGGACAA	CCTCTCTGAG	GGCATTCGCG	
						<u>Rep 68 stop</u>

Fig. 1AU

	2301				2350
42_2	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
42_8	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
42_15	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
42_5b	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
42_1b	AGTGGTGGGA	CTTGAGACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
42_13	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
42_3a	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
42_4	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
42_5a	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
42_10	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
42_3b	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
42_11	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
42_6b	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
43_1	AGTGGTGGGA	CCTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
43_5	AGTGGTGGGA	CCTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
43_12	AGTGGTGGGA	CCTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
43_20	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
43_21	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
43_23	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
43_25	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
44_1	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
44_5	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	AGTGGTGGAA	GCTCAAACCT	GGCCCACCAC	CGCCGAAACC	TAACCAACAA
A3_5	AGTGGTGGAA	GCTCAAACCT	GGCCCACCAC	CGCCGAAACC	TAACCAACAA
A3_7	AGTGGTGGAA	GCTCAAACCT	GGCCCACCAC	CGCCGAAACC	TAACCAACAA
A3_3	AGTGGTGGAA	GCTCAAACCT	GGCCCACCAC	CGCCGAAACC	TAACCAACAA
42_12	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
AAV1	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AGCCCAAAGC	CAACCAGCAA
AAV2	AGTGGTGGAA	GCTCAAACCT	GGCCCACCAC	CACCAAAGCC	CGCAGAGCGG
AAV3	AGTGGTGGGC	TCTGAAACCT	GGAGTCCCTC	AACCCAAAGC	GAACCAACAA
AAV8	AGTGGTGGGC	GCTGAAACCT	GGAGCCCCGA	AGCCCAAAGC	CAACCAGCAA
AAV9	AGTGGTGGGC	GCTGAAACCT	GGAGCCCCGA	AGCCCAAAGC	CAACCAGCAA
AAV7	AGTGGTGGGA	CCTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA
44_2	AGTGGTGGGA	CTTGAAACCT	GGAGCCCCGA	AACCCAAAGC	CAACCAGCAA

Fig. 1AV

	2351				2400
42_2	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
42_8	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
42_15	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
42_5b	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
42_1b	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
42_13	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
42_3a	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
42_4	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
42_5a	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
42_10	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
42_3b	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
42_11	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
42_6b	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
43_1	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
43_5	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
43_12	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
43_20	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
43_21	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
43_23	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
43_25	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
44_1	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
44_5	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	CACCGGGACG	ACAGTAGGGG	TCTTGTGCTT	CCTGGGTACA	AGTACCTCGG
A3_5	CACCGGGACG	ACAGTAGGGG	TCTTGTGCTT	CCTGGGTACA	AGTACCTCGG
A3_7	CACCGGGACG	ACAGTAGGGG	TCTTGTGCTT	CCTGGGTACA	AGTACCTCGG
A3_3	CACCGGGACG	ACAGTAGGGG	TCTTGTGCTT	CCTGGGTACA	AGTACCTCGG
42_12	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
AAV1	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
AAV2	CATAAGGACG	ACAGCAGGGG	TCTTGTGCTT	CCTGGGTACA	AGTACCTCGG
AAV3	CACCAGGACA	ACCGTCGGGG	TCTTGTGCTT	CCGGGTACA	AATACCTCGG
AAV8	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
AAV9	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
AAV7	AAGCAGGACA	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG
44_2	AAGCAGGACG	ACGGCCGGGG	TCTGGTGCTT	CCTGGCTACA	AGTACCTCGG

Fig. 1AW

	2401				2450
42_2	ACCCTTCAAC	GGACTCGACA	AGGGAGAGCC	GGTCAACGAG	GCAGACGCCG
42_8	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
42_15	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
42_5b	ACCCTTCAAC	GGACTCGACA	AGGGAGAGCC	GGTCAACGAG	GCAGACGCCG
42_1b	ACCCTTCAAC	GGACTCGACA	AGGGAGAGCC	GGTCAACGAG	GCAGACGCCG
42_13	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
42_3a	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
42_4	ACCCTTCAAC	GGACTCGACA	AGGGAGAGCC	GGTCAACGAG	GCAGACGCCG
42_5a	ACCCTTCAAC	GGACTCGACA	AGGGAGAGCC	GGTCAACGAG	GCAGACGCCG
42_10	ACCCTTCAAC	GGACTCGACA	AGGGAGAGCC	GGTCAACGAG	GCAGACGCCG
42_3b	ACCCTTCAAC	GGACTCGACA	AGGGAGAGCC	GGTCAACGAG	GCAGACGCCG
42_11	ACCCTTCAAC	GGACTCGACA	AGGGAGAGCC	GGTCAACGCG	GCGGACGCAG
42_6b	ACCCTTCAAC	GGACTCGACA	AGGGAGAGCC	GGTCAACGAG	GCAGACGCCG
43_1	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
43_5	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
43_12	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
43_20	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
43_21	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
43_23	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
43_25	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
44_1	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
44_5	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	ACCCTTCAAC	GGACTCGACA	AAGGAGAGCC	GGTCAACGAG	GCAGACGCCG
A3_5	ACCCTTCAAC	GGACTCGACA	AAGGAGAGCC	GGTCAACGAG	GCAGACGCCG
A3_7	ACCCTTCAAC	GGACTCGACA	AAGGAGAGCC	GGTCAACGAG	GCAGACGCCG
A3_3	ACCCTTCAAC	GGACTCGACA	AAGGAGAGCC	GGTCAACGAG	GCAGACGCCG
42_12	ACCCTTCAAC	GGACTCGACA	AGGGAGAGCC	GGTCAACGAG	GCAGACGCCG
AAV1	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
AAV2	ACCCTTCAAC	GGACTCGACA	AGGGAGAGCC	GGTCAACGAG	GCAGACGCCG
AAV3	ACCCGGTAAC	GGACTCGACA	AAGGAGAGCC	GGTCAACGAG	GCGGACGCGG
AAV8	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
AAV9	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
AAV7	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG
44_2	ACCCTTCAAC	GGACTCGACA	AGGGGGAGCC	CGTCAACGCG	GCGGACGCAG

Fig. 1AX

	2451				2500
42_2	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	AAGCAGCTCG	AGCAGGGGGA
42_8	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
42_15	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
42_5b	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	AAGCAGCTCG	AGCAGGGGGA
42_1b	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	AAGCAGCTCG	AGCAGGGGGA
42_13	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
42_3a	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
42_4	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	AAGCAGCTCG	AGCAGGGGGA
42_5a	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	AAGCAGCTCG	AGCAGGGGGA
42_10	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	AAGCAGCTCG	AGCAGGGGGA
42_3b	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	AAGCAGCTCG	AGCAGGGGGA
42_11	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
42_6b	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	AAGCAGCTCG	AGCAGGGGGA
43_1	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
43_5	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
43_12	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
43_20	CGGCCCTCGA	GCACG.ACAA	AGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
43_21	CGGCCCTCGA	GCACG.ACAA	AGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
43_23	CGGCCCTCGA	GCACG.ACAA	AGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
43_25	CGGCCCTCGA	GCACG.ACAA	AGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
44_1	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
44_5	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
223_10	.....	.....CAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
223_2	.....	.....CAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
223_4	.....	.....CAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
223_5	.....	.....CAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
223_6	.....	.....CAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
223_7	.....	.....CAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
A3_4	CGGCCCTCGA	GCACG.ACAA	AGCCTACGAC	CACCAGCTCA	AGCAAGGGGA
A3_5	CGGCCCTCGA	GCACG.ACAA	AGCCTACGAC	CACCAGCTCA	AGCAAGGGGA
A3_7	CGGCCCTCGA	GCACG.ACAA	AGCCTACGAC	CACCAGCTCA	AGCAAGGGGA
A3_3	CGGCCCTCGA	GCACG.ACAA	AGCCTACGAC	CACCAGCTCA	AGCAAGGGGA
42_12	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	AAGCAGCTCG	AGCAGGGGGA
AAV1	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
AAV2	CGGCCCTCGA	GCACGTACAA	AGCCTACGAC	CGGCAGCTCG	ACAGCGGAGA
AAV3	CAGCCCTCGA	ACACG.ACAA	AGCTTACGAC	CAGCAGCTCA	AGGCCGGTGA
AAV8	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTGC	AGGCCGGTGA
AAV9	CGGCCCTCGA	GCACG.GCAA	GGCCTACGAC	CAGCAGCTGC	AGGCCGGTGA
AAV7	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA
44_2	CGGCCCTCGA	GCACG.ACAA	GGCCTACGAC	CAGCAGCTCA	AAGCGGGTGA

Fig. 1AY

	2501				2550
42_2	CAACCCGTAC	CTCAAGTACA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
42_8	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
42_15	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
42_5b	CAACCCGTAC	CTCAAGTACA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
42_1b	CAACCCGTAC	CTCAAGTACA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
42_13	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
42_3a	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
42_4	CAACCCGTAC	CTCAAGTACA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
42_5a	CAACCCGTAC	CTCAAGTACA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
42_10	CAACCCGTAC	CTCAAGTACA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
42_3b	CAACCCGTAC	CTCAAGTACA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
42_11	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
42_6b	CAACCCGTAC	CTCAAGTACA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
43_1	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
43_5	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
43_12	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
43_20	CAATCCGTAC	CTGCGGTATA	ATCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
43_21	CAATCCGTAC	CTGCGGTATA	ATCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
43_23	CAATCCGTAC	CTGCGGTATA	ATCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
43_25	CAATCCGTAC	CTGCGGTATA	ATCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
44_1	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
44_5	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
223_10	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
223_2	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGTGTC
223_4	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
223_5	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
223_6	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
223_7	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
A3_4	CAACCCGTAC	CTCAAATACA	ACCACGCGGA	CGCTGAATTT	CAGGAGCGTC
A3_5	CAACCCGTAC	CTCAAATACA	ACCACGCGGA	CGCTGAATTT	CAGGAGCGTC
A3_7	CAACCCGTAC	CTCAAATACA	ACCACGCGGA	CGCTGAATTT	CAGGAGCGTC
A3_3	CAACCCGTAC	CTCAAATACA	ACCACGCGGA	CGCTGAATTT	CAGGAGCGTC
42_12	CAACCCGTAC	CTCAAGTACA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
AAV1	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
AAV2	CAACCCGTAC	CTCAAGTACA	ACCACGCCGA	CGCGGAGTTT	CAGGAGCGCC
AAV3	CAACCCGTAC	CTCAAGTACA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
AAV8	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
AAV9	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
AAV7	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC
44_2	CAATCCGTAC	CTGCGGTATA	ACCACGCCGA	CGCCGAGTTT	CAGGAGCGTC

Fig. 1A2

	2551				2600
42_2	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_8	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_15	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_5b	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_1b	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_13	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_3a	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_4	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_5a	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_10	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_3b	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_11	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_6b	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
43_1	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
43_5	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
43_12	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
43_20	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
43_21	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
43_23	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
43_25	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
44_1	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
44_5	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
223_10	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
223_2	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
223_4	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
223_5	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
223_6	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
223_7	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
A3_4	TTCAAGAAGA	TACGTCTTTC	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
A3_5	TTCAAGAAGA	TACGTCTTTC	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
A3_7	TTCAAGAAGA	TACGTCTTTC	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
A3_3	TTCAAGAAGA	TACGTCTTTC	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
42_12	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
AAV1	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
AAV2	TTAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGACGAGC	AGTCTTCCAG
AAV3	TTCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TTGGCAGAGC	AGTCTTCCAG
AAV8	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
AAV9	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
AAV7	TGCAAGAAGA	TACGTCAATT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG
44_2	TGCAAGAAGA	TACGTCTTTT	GGGGGCAACC	TCGGGCGAGC	AGTCTTCCAG

Fig. 1AAA

	2601				2650
42_2	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
42_8	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
42_15	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
42_5b	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
42_1b	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
42_13	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
42_3a	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
42_4	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
42_5a	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
42_10	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
42_3b	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
42_11	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
42_6b	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
43_1	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
43_5	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
43_12	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
43_20	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
43_21	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
43_23	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
43_25	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
44_1	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
44_5	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
223_10	GCCAAAAAGC	GGGTTCTCGA	ACCTCTTGGT	CTGGTTGAGA	CGCCAGCTAA
223_2	GCCAAAAAGC	GGGTTCTCGA	ACCTCTTGGT	CTGGTTGAGA	CGCCAGCTAA
223_4	GCCAAAAAGC	GGGTTCTCGA	ACCTCTTGGT	CTGGTTGAGA	CGCCAGCTAA
223_5	GCCAAAAAGC	GGGTTCTCGA	ACCTCTTGGT	CTGGTTGAGA	CGCCAGCTAA
223_6	GCCAAAAAGC	GGGTTCTCGA	ACCTCTTGGT	CTGGTTGAGA	CGCCAGCTAA
223_7	GCCAAAAAGC	GGGTTCTCGA	ACCTCTTGGT	CTGGTTGAGA	CGCCAGCTAA
A3_4	GCCAAAAAGA	GGGTACTCGA	GCCTCTTGGT	CTGGTTGAGG	AAGCTGTTAA
A3_5	GCCAAAAAGA	GGGTACTCGA	GCCTCTTGGT	CTGGTTGAGG	AAGCTGTTAA
A3_7	GCCAAAAAGA	GGGTACTCGA	GCCTCTTGGT	CTGGTTGAGG	AAGCTGTTAA
A3_3	GCCAAAAAGA	GGGTACTCGA	GCCTCTTGGT	CTGGTTGAGG	AAGCTGTTAA
42_12	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
AAV1	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
AAV2	GCGAAAAAGA	GGGTTCTTGA	ACCTCTGGGC	CTGGTTGAGG	AACCTGTTAA
AAV3	GCCAAAAAGA	GGATCCTTGA	GCCTCTTGGT	CTGGTTGAGG	AAGCAGCTAA
AAV8	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
AAV9	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
AAV7	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA
44_2	GCCAAGAAGC	GGGTTCTCGA	ACCTCTCGGT	CTGGTTGAGG	AAGGCGCTAA

Fig. 1AAB

2651

2700

vp2 start  
 42\_2 GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC .....  
 42\_8 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC  
 42\_15 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC  
 42\_5b GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC  
 42\_1b GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC .....  
 42\_13 GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC .....  
 42\_3a GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC .....  
 42\_4 GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC .....  
 42\_5a GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC .....  
 42\_10 GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC .....  
 42\_3b GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC .....  
 42\_11 GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC .....  
 42\_6b GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC  
 43\_1 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCT CAGCGTTCCC  
 43\_5 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCT CAGCGTTCCC  
 43\_12 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCT CAGCGTTCCC  
 43\_20 GACGGCTCCT GGAAAGAAGA GACTGGTAGA GCAGTCGCCA CAAGAG...C  
 43\_21 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCAGTCGCCA CAAGAG...C  
 43\_23 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCAGTCGCCA CAAGAG...C  
 43\_25 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCAGTCGCCA CAAGAG...C  
 44\_1 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC  
 44\_5 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC  
 223\_10 GACGGCACCT GGAAAGAAGC GACCGGTAGA CTCGCCA... ..  
 223\_2 GACGGCACCT GGAAAGAAGC GACCGGTAGA CTCGCCA... ..  
 223\_4 GACGGCACCT GGAAAGAAGC GACCGGTAGA CTCGCCA... ..  
 223\_5 GACGGCACCT GGAAAGAAGC GACCGGTAGA CTCGCCA... ..  
 223\_6 GACGGCACCT GGAAAGAAGC GACCGGTAGA CTCGCCA... ..  
 223\_7 GACGGCACCT GGAAAGAAGC GACCGGTAGA CTCGCCA... ..  
 A3\_4 GACGGCTCCT GGAAAAAAGA GACCTATAGA GCAGTCTCCT GCAGAA...C  
 A3\_5 GACGGCTCCT GGAAAAAAGA GACCTATAGA GCAGTCTCCT GCAGAA...C  
 A3\_7 GACGGCTCCT GGAAAAAAGA GACCTATAGA GCAGTCTCCT GCAGAA...C  
 A3\_3 GACGGCTCCT GGAAAAAAGA GACCTATAGA GCAGTCTCCT GCAGAA...C  
 42\_12 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC  
 AAV1 GACGGCTCCT GGAAAGAAAC GTCCGGTAGA GCAGTCGCCA CAAGAG...C  
 AAV2 GACGGCTCCG GGAAAAAAGA GGCCGGTAGA GCACTCTCCT GTGGAG...C  
 AAV3 AACGGCTCCT GGAAAGAAGG GGGCTGTAGA TCAGTCTCCT CAGGAA...C  
 AAV8 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC  
 AAV9 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC  
 AAV7 GACGGCTCCT GCAAAGAAGA GACCGGTAGA GCCGTACCT CAGCGTTCCC  
 44\_2 GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC  
 vp2 start

Fig. 1AAC

	2701				2750
42_2	..GACTCCTC	CACGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCTAAAAAG
42_8	CAGACTCCTC	TACGGGCATC	GGCAAGACAG	GCCAGCAGCC	CGCGAAAAAG
42_15	CAGACTCCTC	TACGGGCATC	GGCAAGACAG	GCCAGCAGCC	CGCGAAAAAG
42_5b	CAGACTCCTC	TACGGGCATC	GGCAAGACAG	GCCAGCAGCC	CGCGAAAAAG
42_1b	..GACTCCTC	CACGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCTAAAAAG
42_13	..GACTCCTC	CACGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCTAAAAAG
42_3a	..GACTCCTC	CACGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCTAAAAAG
42_4	..GACTCCTC	CACGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCTAAAAAG
42_5a	..GACTCCTC	CACGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCTAAAAAG
42_10	..GACTCCTC	CACGGGCATC	GGCAGGAAAG	GCCAGCAGCC	CGCTAAAAAG
42_3b	..GACTCCTC	CACGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCTAAAAAG
42_11	..GACTCCTC	CACGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCTAAAAAG
42_6b	CAGACTCCTC	TACGGGCATC	GGCAAGACAG	GCCAGCAGCC	CGCGAAAAAG
43_1	CCGACTCCTC	CACGGGCATC	GGCAAGAAAG	GCCACCAGCC	CGCGAGAAAG
43_5	CCGACTCCTC	CACGGGCATC	GGCAAGAAAG	GCCACCAGCC	CGCGAGAAAG
43_12	CCGACTCCTC	CACGGGCATC	GGCAAGAAAG	GCCACCAGCC	CGCGAGAAAG
43_20	CAGACTCCTC	CTCGGGCATC	GGCAAGACAG	GCCAGCAGCC	CGCTAAAAAG
43_21	CAGACTCCTC	CTCGGGCATC	GGCAAGACAG	GCCAGCAGCC	CGCTAAAAAG
43_23	CAGACTCCTC	CTCGGGCATC	GGCAAGACAG	GCCAGCAGCC	CGCTAAAAAG
43_25	CAGACTCCTC	CTCGGGCATC	GGCAAGACAG	GCCAGCAGCC	CGCTAAAAAG
44_1	CAGACTCCTC	TACGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCGAAAAAG
44_5	CAGACTCCTC	TACGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCGAAAAAG
223_10	..GACTCCAC	CTCGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCGAAAAAG
223_2	..GACTCCAC	CTCGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCGAAAAAG
223_4	..GACTCCAC	CTCGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCGAAAAAG
223_5	..GACTCCAC	CTCGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCGAAAAAG
223_6	..GACTCCAC	CTCGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCGAAAAAG
223_7	..GACTCCAC	CTCGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCGAAAAAG
A3_4	CGGACTCTTC	CTCGGGCATC	GGCAATCAG	GCCAGCAGCC	CGCTAAGAAA
A3_5	CGGACTCTTC	CTCGGGCATC	GGCAATCAG	GCCAGCAGCC	CGCTAAGAAA
A3_7	CGGACTCTTC	CTCGGGCATC	GGCAATCAG	GCCAGCAGCC	CGCTAAGAAA
A3_3	CGGACTCTTC	CTCGGGCATC	GGCAATCAG	GCCAGCAGCC	CGCTAAGAAA
42_12	CAGACTCCTC	TACGGGCATC	GGCAAGACAG	GCCAGCAGCC	CGCGAAAAAG
AAV1	CAGACTCCTC	CTCGGGCATC	GGCAAGACAG	GCCAGCAGCC	CGCTAAAAAG
AAV2	CAGACTCCTC	CTCGGGAACC	GGAAAGGCGG	GCCAGCAGCC	TGCAAGAAAA
AAV3	CGGACTCATC	ATCTGGTGTT	GGCAAATCGG	GCAAACAGCC	TGCCAGAAAA
AAV8	CAGACTCCTC	TACGGGCATC	GGCAAGAAAG	GCCAACAGCC	CGCCAGAAAA
AAV9	CAGACTCCTC	TACGGGCATC	GGCAAGAAAG	GCCAACAGCC	CGCCAGAAAA
AAV7	CCGACTCCTC	CACGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCCAGAAAG
44_2	CAGACTCCTC	TACGGGCATC	GGCAAGAAAG	GCCAGCAGCC	CGCGAAAAAG

Fig. 1AAD

	2751					2800
42_2	AAGCTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCCCA	
42_8	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
42_15	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
42_5b	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
42_1b	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
42_13	AAGCTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
42_3a	AAGCTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
42_4	AAGCTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
42_5a	AAGCTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCCCA	
42_10	AAGCTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
42_3b	AAGCTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
42_11	AAGCTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
42_6b	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
43_1	AGACTGAACT	TTGGGCAGAC	TGGCGACTCG	GAGTCAGTCC	CCGACCCTCA	
43_5	AGACTGAACT	TTGGGCAGAC	TGGCGACTCG	GAGTCAGTCC	CCGACCCTCA	
43_12	AGACTGAACT	TTGGGCAGAC	TGGCGACTCG	GAGTCAGTCC	CCGACCCTCA	
43_20	AGACTCAATT	TTGGTCAGAC	TGGCGACTCA	GAGTCAGTCC	CCGACCCACA	
43_21	AGACTCAATT	TTGGTCAGAC	TGGCGACTCA	GAGTCAGTCC	CCGACCCACA	
43_23	AGACTCAATT	TTGGTCAGAC	TGGCGACTCA	GAGTCAGTCC	CCGACCCACA	
43_25	AGACTCAATT	TTGGTCAGAC	TGGCGACTCA	GAGTCAGTCC	CCGACCCACA	
44_1	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
44_5	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
223_10	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTCC	CCGACCCTCA	
223_2	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTCC	CCGACCCTCA	
223_4	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGCCAGTCC	CCGACCCTCA	
223_5	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGCCAGTCC	CCGACCCTCA	
223_6	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTCC	CCGACCCTCA	
223_7	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTCC	CCGACCCTCA	
A3_4	AGACTCAATT	TTGGTCAGAC	TGGCGACACA	GAGTCAGTCC	CAGACCCTCA	
A3_5	AGACTCAATT	TTGGTCAGAC	TGGCGACACA	GAGTCAGTCC	CAGACCCTCA	
A3_7	AGACTCAATT	TTGGTCAGAC	TGGCGACACA	GAGTCAGTCC	CAGACCCTCA	
A3_3	AGACTCAATT	TTGGTCAGAC	TGGCGACACA	GAGTCAGTCC	CAGGCCCTCA	
42_12	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	
AAV1	AGACTCAATT	TTGGTCAGAC	TGGCGACTCA	GAGTCAGTCC	CCGATCCACA	
AAV2	AGATTGAATT	TTGGTCAGAC	TGGAGACGCA	GAATCAGTAC	CTGACCCCCA	
AAV3	AGACTAAATT	TCGGTCAGAC	TGGAGACTCA	GAGTCAGTCC	CAGACCCTCA	
AAV8	AGACTCAATT	TTGGTCAGAC	TGGCGACTCA	GAGTCAGTTC	CAGACCCTCA	
AAV9	AGACTCAATT	TTGGTCAGAC	TGGCGACTCA	GAGTCAGTTC	CAGACCCTCA	
AAV7	AGACTCAATT	TCGGTCAGAC	TGGCGACTCA	GAGTCAGTCC	CCGACCCTCA	
44_2	AGACTCAACT	TTGGGCAGAC	TGGCGACTCA	GAGTCAGTGC	CCGACCCTCA	

Fig. 1AAE

	2801				2850	
					vp3	start
42_2	ACCTCTCGGA	GAACCTCCCG	CCGCGCCCTC	AGGTCTGGGA	TCTGGTACAA	
42_8	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
42_15	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
42_5b	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
42_1b	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGCACAA	
42_13	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
42_3a	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
42_4	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
42_5a	ACCTCTCGGA	GAACCTCCCG	CCGCGCCCTC	AGGTCTGGGA	TCTGGTACAA	
42_10	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
42_3b	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
42_11	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
42_6b	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
43_1	ACCAATCGGA	GAACCACCAG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
43_5	ACCAATCGGA	GAACCACCAG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
43_12	ACCAATCGGA	GAACCACCAG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
43_20	ACCTCTCGGA	GAACCTCCAG	CAGGCCCTC	AGGTCTGGGA	CCTAATACAA	
43_21	ACCTCTCGGA	GAACCTCCAG	CAGGCCCTC	AGGTCTGGGA	CCTAATACAA	
43_23	ACCTCTCGGA	GAACCTCCAG	CAGGCCCTC	AGGTCTGGGA	CCTAATACAA	
43_25	ACCTCTCGGA	GAACCTCCAG	CAGGCCCTC	AGGTCTGGGA	CCTAATACAA	
44_1	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
44_5	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
223_10	ACCAATCGGA	GAACCACCAG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
223_2	ACCAATCGGA	GAACCACCAG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
223_4	ACCAATCGGA	GAACCACCAG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
223_5	ACCAATCGGA	GAACCACCAG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
223_6	ACCAATCGGA	GAACCACCAG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
223_7	ACCAATCGGA	GAACCACCAG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
A3_4	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTGTGGGA	TCTAATACAA	
A3_5	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTGTGGGA	TCTAATACAA	
A3_7	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTGTGGGA	TCTAATACAA	
A3_3	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTGTGGGA	TCTAATACAA	
42_12	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
AAV1	ACCTCTCGGA	GAACCTCCAG	CAACCCCCGC	TGCTGTGGGA	CCTACTACAA	
AAV2	GCCTCTCGGA	CAGCCACCAG	CAGGCCCTC	TGGTCTGGGA	ACTAATACGA	
AAV3	ACCTCTCGGA	GAACCACCAG	CAGCCCCAC	AAGTTTGGGA	TCTAATACAA	
AAV8	ACCTCTCGGA	GAACCTCCAG	CAGCGCCCTC	TGGTGTGGGA	CCTAATACAA	
AAV9	ACCTCTCGGA	GAACCTCCAG	CAGCGCCCTC	TGGTGTGGGA	CCTAATACAA	
AAV7	ACCTCTCGGA	GAACCTCCAG	CAGCGCCCTC	TAGTGTGGGA	TCTGGTACAG	
44_2	ACCAATCGGA	GAACCCCCCG	CAGGCCCTC	TGGTCTGGGA	TCTGGTACAA	
					vp3	start

Fig. 1AAF

2851 2900

←vp3 start codon

42_2	TGGCTGCAGG	CGGTGGCGCA	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_8	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_15	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_5b	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_1b	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_13	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_3a	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_4	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_5a	TGGCTGCAGG	CGGTGGCGCA	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_10	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_3b	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_11	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_6b	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
43_1	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
43_5	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
43_12	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
43_20	TGGCTTCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
43_21	TGGCTTCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
43_23	TGGCTTCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
43_25	TGGCTTCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
44_1	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
44_5	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
223_10	TGGCTGCAGG	CGGTGGCGCA	CCAATGGCTG	ACAATAACGA	GGGCGCCGAC
223_2	TGGTTGCAGG	CGGTGGCGCA	CCAATGGCTG	ACAATAACGA	GGGCGCCGAC
223_4	TGGCTGCAGG	CGGTGGCGCA	CCAATGGCTG	ACAATAACGA	GGGCGCCGAC
223_5	TGGCTGCAGG	CGGTGGCGCA	CCAATGGCTG	ACAATAACGA	GGGCGCCGAC
223_6	TGGCTGCAGG	CGGTGGCGCA	CCAATGGCTG	ACAATAGCGA	GGGCGCCGAC
223_7	TGGCTGCAGG	CGGTGGCGCA	CCAATGGCTG	ACAATAACGA	GGGCGCCGAC
A3_4	TGGCTTCAGG	CGGTGGGGCA	CCAATGGCAG	ACGATAACGA	AGGCGCCGAC
A3_5	TGGCTTCAGG	CGGTGGGGCA	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
A3_7	TGGCTTCAGG	CGGTGGGGCA	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
A3_3	TGGCTTCAGG	CGGTGGGGCA	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
42_12	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
AAV1	TGGCTTCAGG	CGGTGGCGCA	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
AAV2	TGGCTACAGG	CAGTGGCGCA	CCAATGGCAG	ACAATAACGA	GGGCGCCGAC
AAV3	TGGCTTCAGG	CGGTGGCGCA	CCAATGGCAG	ACAATAACGA	GGGTGCCGAT
AAV8	TGGCTGCAGG	CGGTGGCGCA	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
AAV9	TGGCTGCAGG	CGGTGGCGCA	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC
AAV7	TGGCTGCAGG	CGGTGGCGCA	CCAATGGCAG	ACAATAACGA	AGGTGCCGAC
44_2	TGGCTGCAGG	CGGTGGCGCT	CCAATGGCAG	ACAATAACGA	AGGCGCCGAC

vp3 start codon (cont'd)

Fig. 1AAG

	2901				2950
42_2	GGAGTGGGTA	ATGCCTCCGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
42_8	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
42_15	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
42_5b	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
42_1b	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
42_13	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
42_3a	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATAGCTGGG
42_4	GGAGTGGGTA	ATGCCTCCGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
42_5a	GGAGTGGGTA	ATGCCTCCGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
42_10	GGAGTGGGTA	ATGCCTCCGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
42_3b	GGAGTGGGTA	ATGCCTCCGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
42_11	GGAGTGGGTA	ATGCCTCCGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
42_6b	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
43_1	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
43_5	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
43_12	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
43_20	GGAGTGGGTA	ATTCCCTCGGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
43_21	GGAGTGGGTA	ATTCCCTCGGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
43_23	GGAGTGGGTA	ATTCCCTCGGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
43_25	GGAGTGGGTA	ATTCCCTCGGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
44_1	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
44_5	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
223_10	GGAGTGGGTA	ATGCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
223_2	GGAGTGGGTA	ATGCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
223_4	GGAGTGGGTA	ATGCCTCAGG	AAATTGGCAT	TGCGATTCCA	CACGGCTGGG
223_5	GGAGTGGGTA	ATGCCTCAGG	AAATTGGCAT	TGCGATTCCA	CACGGCTGGG
223_6	GGAGTGGGTA	ATGCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
223_7	GGAGTGGGTA	ATGCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
A3_4	GGAGTGGGTA	ATTCCCTCGGG	AAATTGGCAT	TGCGATTCCA	CATGGATGGG
A3_5	GGAGTGGGTA	ATTCCCTCGGG	AAATTGGCAT	TGCGATTCCA	CATGGATGGG
A3_7	GGAGTGGGTA	ATTCCCTCGGG	AAATTGGCAT	TGCGATTCCA	CATGGATGGG
A3_3	GGAGTGGGTA	ATTCCCTCGGG	AAATTGGCAT	TGCGATTCCA	CATGGATGGG
42_12	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
AAV1	GGAGTGGGTA	ATGCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
AAV2	GGAGTGGGTA	ATTCCCTCCGG	AAATTGGCAT	TGCGATTCCA	CATGGATGGG
AAV3	GGAGTGGGTA	ATTCCCTCAGG	AAATTGGCAT	TGCGATTCCC	AATGGCTGGG
AAV8	GGAGTGGGTA	GTTCCCTCGGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
AAV9	GGAGTGGGTA	ATTCCCTCGGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
AAV7	GGAGTGGGTA	ATGCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
AAV10	GGTA	ATTCCCTCCGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
AAV11	GGTA	ATTCCCTCCGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
AAV12	GGTA	ATTCCCTCCGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG
44_2	GGAGTGGGTA	GTTCCCTCAGG	AAATTGGCAT	TGCGATTCCA	CATGGCTGGG

FIG. 1AAH

	2951				3000
42_2	CGACAGAGTC	ATCACCACCA	GCACCCGCAC	CTGGGGCCCTG	CCCACCTACA
42_8	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTC	CCCACCTACA
42_15	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTC	CCCACCTACA
42_5b	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTC	CCCACCTACA
42_1b	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTC	CCCACCTACA
42_13	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTC	CCCACCTACA
42_3a	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTC	CCCACCTACA
42_4	CGACAGAGTC	ATCACCACCA	GCACCCGCAC	CTGGGGCCCTG	CCCACCTACA
42_5a	CGACAGAGTC	ATCACCACCA	GCACCCGCAC	CTGGGGCCCTG	CCCACCTACA
42_10	CGACAGAGTC	ATCACCACCA	GCACCCGCAC	CTGGGGCCCTG	CCCACCTACA
42_3b	CGACAGAGTC	ATCACCACCA	GCACCCGCAC	CTGGGGCCCTG	CCCACCTACA
42_11	CGACAGAGTC	ATCACCACCA	GCACCCGCAC	CTGGGGCCCTG	CCCACCTACA
42_6b	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTC	CCCACCTACA
43_1	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
43_5	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
43_12	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
43_20	GGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
43_21	GGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
43_23	GGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
43_25	GGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
44_1	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTC	CCCACCTACA
44_5	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTC	CCCACCTACA
223_10	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
223_2	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
223_4	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
223_5	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
223_6	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
223_7	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
A3_4	CGACAGAGTT	ATCACCACCA	GCACAAGAAC	CTGGGGCCCTC	CCCACCTACA
A3_5	CGACAGAGTT	ATCACCACCA	GCACAAGAAC	CTGGGGCCCTC	CCCACCTACA
A3_7	CGACAGAGTT	ATCACCACCA	GCACAAGAAC	CTGGGGCCCTC	CCCACCTACA
A3_3	CGACAGAGTT	ATCACCACCA	GCACAAGAAC	CTGGGGCCCTC	CCCACCTACA
42_12	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTC	CCCACCTACA
AAV1	CGACAGAGTC	ATCACCACCA	GCACCCGCAC	CTGGGGCCCTG	CCCACCTACA
AAV2	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
AAV3	CGACAGAGTC	ATCACCACCA	GCACCAGAAC	CTGGGGCCCTG	CCCACCTACA
AAV8	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
AAV9	GGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCATTG	CCCACCTACA
AAV7	CGACAGAGTC	ATTACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCCACCTACA
AAV10	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGTCCTG	CCCACCTACA
AAV11	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTG	CCAACCTACA
AAV12	CGACAGAGTC	ATTACCACCA	GCACCCGGAC	CTGGGGCCCTG	CCCACCTACA
44_2	CGACAGAGTC	ATCACCACCA	GCACCCGAAC	CTGGGGCCCTC	CCCACCTACA

Fig. 1AAI

```

3001
42_2 ACAACCACCT CTACAAGCAG ATATCAA..G TCAGAGCGGG GCT....ACC
42_8 ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
42_15 ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
42_5b ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
42_1b ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
42_13 ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
42_3a ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
42_4 ACAACCACCT CTACAAGCAG ATATCAA... .GTCAGAGCG GGGC..TACC
42_5a ACAACCACCT CTACAAGCAG ATATCAA... .GTCAGAGCG GGGC..TACC
42_10 ACAACCACCT CTACAAGCAG ATATCAA..G TCAGAGCGGG GCTA....CC
42_3b ACAACCACCT CTACAAGCAG ATATCAA..G TCAGAGCGGG GCTA....CC
42_11 ACAACCACCT CTACAAGCAG ATATCAA..G TCAGAGCGGG GCTA....CC
42_6b ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
43_1 ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACT
43_5 ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACT
43_12 ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACT
43_20 ACAACCACCT CTACAAGCAA ATCTCCAACG GCACCTCGGG AGGAAGCACC
43_21 ACAACCACCT CTACAAGCAA ATCTCCAACG GCACCTCGGG AGGAAGCACC
43_23 ACAACCACCT CTACAAGCAA ATCTCCAACG GCACCTCGGG AGGAAGCACC
43_25 ACAACCACCT CTACAAGCAA ATCTCCAACG GCACCTCGGG AGGAAGCACC
44_1 ACAACCACCT CTACAAGCAA ATCTCCAACG GGACTTCGGG AGGAAGCACC
44_5 ACAACCACCT CTACAAGCAA ATCTCCAACG GGACTTCGGG AGGAAGCACC
223_10 ACAACCACCT CTACAAGCAA ATCTCCAGTC AGTCAGCAGG GAG...CACC
223_2 ACAACCACCT CTACAAGCAA ATCTCCAGTC AGTCAGCAGG GAG...CACC
223_4 ACAACCACCT CTACAAGCAA ATCTCCAGTC AGTCAGCAGG GAG...CACC
223_5 ACAACCACCT CTACAAGCAA ATCTCCAGTC AGTCAGCAGG GAG...CACC
223_6 ACAACCACCT CTACAAGCAA ATCTCCAGTC AGTCAGCAGG GAG...CACC
223_7 ACAACCACCT CTACAAGCAA ATCTCCAGTC AGTCAGCAGG GAG...CACC
A3_4 ATAATCACCT CTACAAGCAA ATCTCCA... GCGAATCGGG AGC...CACC
A3_5 ATAATCACCT CTACAAGCAA ATCTCCA... GCGAATCGGG AGC...CACC
A3_7 ATAATCGCCT CTACAAGCAA ATCTCCA... GCGAATCGGG AGC...CACC
A3_3 ATAATCACCT CTACAAGCAA ATCTCCA... GCGAATCGGG AGC...CACC
42_12 ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
AAV1 ATAACCACCT CTACAAGCAA ATCTCCAGTG CTTCAACGGG .GG..CCAGC
AAV2 ACAACCACCT CTACAAACAA ATTTCCA... GCCAATCAGG AGC...CTCG
AAV3 ACAACCACCT CTACAAGCAA ATCTCCA... GCCAATCAGG AGC...TTCA
AAV8 ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAGCCACC
AAV9 ACAACCACCT CTACAAGCAA ATCTCCAATG GAACATCGGG AGGAAGCACC
AAV7 ACAACCACCT CTACAAGCAA ATCTCCAGTG AAAGTGCAGG TAG...TACC
AAV10 ACAACCACAT CTACAAGCAA ATCTCCAGCG AGACAGGAGC CACCAACGAC
AAV11 ACAACCACCT CTACAAACAA ATCTCCAGCG CTTCAACGGG GGCCAGCAAC
AAV12 ACAACCACCT CTACAAGCAA ATCTCCAGCC AATCGGGTGC CACCAACGAC
44_2 ACAACCACCT CTACAAGCAA ATCTCCAACG GGACTTCGGG AGGAAGCACC

```

Fig. 1AAJ

	3051					3100
42_2	AACGACAACC	ACTTCTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_8	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_15	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_5b	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_1b	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_13	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_3a	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_4	AACGACAACC	ACTTCTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_5a	AACGACAACC	ACTTCTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_10	AACGACAACC	ACTTCTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_3b	AACGACAACC	ACTTCTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_11	AACGACAACC	ACTTCTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_6b	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
43_1	AACGACAACA	CCTACTTTGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
43_5	AACGACAACA	CCTACTTTGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
43_12	AACGACAACA	CCTACTTTGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
43_20	AACGACAACA	CCTATTTTGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
43_21	AACGACAACA	CCTATTTTGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
43_23	AACGACAACA	CCTATTTTGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
43_25	AACGACAACA	CCTATTTTGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
44_1	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
44_5	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
223_10	AACGATAACG	TCTATTTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
223_2	AACGATAACG	TCTATTTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
223_4	AACGATAACG	TCTATTTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
223_5	AACGATAACG	TCTATTTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
223_6	AACGATAACG	TCTATTTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
223_7	AACGATAACG	TCTATTTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
A3_4	AACGACAACC	ACTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
A3_5	AACGACAACC	ACTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
A3_7	AACGACAACC	ACTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
A3_3	AACGACAACC	ACTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
42_12	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
AAV1	AACGACAACC	ACTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
AAV2	AACGACAATC	ACTACTTTGG	CTACAGCACC	CCTTGGGGGT	ATTTTGACTT	
AAV3	AACGACAACC	ACTACTTTGG	CTACAGCACC	CCTTGGGGGT	ATTTTGACTT	
AAV8	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
AAV9	AACGACAACA	CCTACTTTGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
AAV7	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
AAV10	AACCACTACT	TCGGCTACAG	C.....ACC	CCCTGGGGGT	ATTTTGACTT	
AAV11	...GACAACC	ACTACTTTGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	
AAV12	AACCACTACT	TCGGCTA...	...CAGCACC	CCTTGGGGGT	ATTTTGACTT	
44_2	AACGACAACA	CCTACTTCGG	CTACAGCACC	CCCTGGGGGT	ATTTTGACTT	

Fig. 1AAK

	3101				3150
42_2	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
42_8	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
42_15	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
42_5b	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
42_1b	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
42_13	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
42_3a	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
42_4	CAACAGATTC	CACTGCCACT	TCTCATCACG	TGACTGGCAG	CGACTCATCA
42_5a	CAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
42_10	CAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
42_3b	CAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
42_11	CAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
42_6b	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
43_1	CAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
43_5	CAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
43_12	CAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
43_20	CAACAGATTC	CACTGTCACT	TTTCACCACG	TGACTGGCAA	CGACTCATCA
43_21	CAACAGATTC	CACTGTCACT	TTTCACCACG	TGACTGGCAA	CGACTCATCA
43_23	CAACAGATTC	CACTGTCACT	TTTCACCACG	TGACTGGCAA	CGACTCATCA
43_25	CAACAGATTC	CACTGTCACT	TTTCACCACG	TGACTGGCAA	CGACTCATCA
44_1	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
44_5	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
223_10	CAACAGATTC	CATTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTTATCA
223_2	CAACAGATTC	CATTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTTATCA
223_4	CAACAGATTC	CATTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTTATCA
223_5	CAACAGATTC	CATTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTTATCA
223_6	CAACAGATTC	CATTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTTATCA
223_7	CAACAGATTC	CATTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTTATCA
A3_4	TAACAGATTC	CACTGTCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
A3_5	TAACAGATTC	CACTGTCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
A3_7	TAACAGATTC	CACTGTCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
A3_3	TAACAGATTC	CACTGTCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
42_12	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
AAV1	CAACAGATTC	CACTGCCACT	TTTCACCACG	TGACTGGCAG	CGACTCATCA
AAV2	CAACAGATTC	CACTGCCACT	TTTCACCACG	TGACTGGCAA	AGACTCATCA
AAV3	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATTA
AAV8	TAACAGATTC	CACTGCCACT	TTTCACCACG	TGACTGGCAG	CGACTCATCA
AAV9	CAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
AAV7	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
AAV10	TAACAGATTC	CACTGCCACT	TTTCACCACG	TGACTGGCAG	CGACTCATCA
AAV11	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
AAV12	CAACAGATTC	CACTGCCATT	TCTCACCACG	TGACTGGCAG	CGACTCATCA
44_2	TAACAGATTC	CACTGCCACT	TCTCACCACG	TGACTGGCAG	CGACTCATCA

Fig. 1AAL

	3151				3200		
42_2	ACAACAAC	TGG	GATTCC	GG	CCCAGAAAGC	TGCGGTTCAA	GTTGTTCAAC
42_8	ACAACAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC
42_15	ACAACAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC
42_5b	ACAACAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC
42_1b	ACAACAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC
42_13	ACAACAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC
42_3a	ACAACAGCTG	GGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC
42_4	ACAACAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC
42_5a	ACAACAACCG	GGG	GATTCC	GG	CCCAGAAAGC	TGCGGTTCAA	GTTGTTCAAC
42_10	ACAACAAC	TGG	GATTCC	GG	CCCAGAAAGC	TGCGGTTCAA	GTTGTTCAAC
42_3b	ACAACAAC	TGG	GATTCC	GG	CCCAGAAAGC	TGCGGTTCAA	GTTGTTCAAC
42_11	ACAACAAC	TGG	GATTCC	GG	CCCAGAAAGC	TGCGGTTCAA	GTTGTTCAAC
42_6b	ACAACAAC	TGG	GATTCC	GG	CCCAGAAAGC	TGCGGTTCAA	GTTGTTCAAC
43_1	ACAATAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC
43_5	ACAATAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC
43_12	ACAATAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC
43_20	ACAACAATTG	GGG	GATTCC	GG	CCCAAAAGAC	TCAACTTCAA	GCTGTTCAAC
43_21	ACAACAATTG	GGG	GATTCC	GG	CCCAAAAGAC	TCAACTTCAA	GCTGTTCAAC
43_23	ACAACAATTG	GGG	GATTCC	GG	CCCAAAAGAC	TCAACTTCAA	GCTGTTCAAC
43_25	ACAACAATTG	GGG	GATTCC	GG	CCCAAAAGAC	TCAACTTCAA	GCTGTTCAAC
44_1	ACAACAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC
44_5	ACAACAAC	TGG	GATTCC	GG	CCCAAGAGAC	CCAACTTCAA	GCTCTTCAAC
223_10	ACAACAAC	TGG	GATTCC	GG	CCCAAGAAGC	TCAACTTCAA	GCTCTTCAAC
223_2	ACAACAAC	TGG	GATTCC	GG	CCCAAGAAGC	TCAACTTCAA	GCTCTTCAAC
223_4	ACAACAAC	TGG	GATTCC	GG	CCCAAGAAGC	TCAACTTCAA	GCTCTTCAAC
223_5	ACAACAAC	TGG	GATTCC	GG	CCCAAGAAGC	TCAACTTCAA	GCTCTTCAAC
223_6	ACAACAAC	TGG	GATTCC	GG	CCCAAGAAGC	TCAACTTCAA	GCTCTTCAAC
223_7	ACAACAAC	TGG	GATTCC	GG	CCCAAGAAGC	TCAACTTCAA	GCTCTTCAAC
A3_4	ACAACAAC	TGG	GATTTAGA	CCCAAGAAAC	TCAATTTCAA	GCTCTTCAAC	
A3_5	ATAACAAC	TGG	GATTTAGA	CCCAAGAAAC	TCAATTTCAA	GCTCTTCAAC	
A3_7	ACAACAAC	TGG	GATTTAGA	CCCAAGAAAC	TCAATTTCAA	GCTCTTCAAC	
A3_3	ACAACAAC	TGG	GATTTAGA	CCCAAGAAAC	TCAATTTCAA	GCTCTTCAAC	
42_12	ACAACAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC
AAV1	ACAACAATTG	GGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	ACTCTTCAAC
AAV2	ACAACAAC	TGG	GATTCCGA	CCCAAGAGAC	TCAACTTCAA	GCTCTTTAAC	
AAV3	ACAACAAC	TGG	GATTCC	GG	CCCAAGAAAC	TCAGCTTCAA	GCTCTTCAAC
AAV8	ACAACAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAGCTTCAA	GCTCTTCAAC
AAV9	ACAACAAC	TGG	GATTCC	GG	CCAAAGAGAC	TCAACTTCAA	GCTGTTCAAC
AAV7	ACAACAAC	TGG	GATTCC	GG	CCCAAGAAGC	TGCGGTTCAA	GCTCTTCAAC
AAV10	ACAACAAC	TGG	GATTC				
AAV11	ACAACAAC	TGG	GATTC				
AAV12	ACAACAAC	TGG	GATTC				
44_2	ACAACAAC	TGG	GATTCC	GG	CCCAAGAGAC	TCAACTTCAA	GCTCTTCAAC

Fig. 1AAM

	3201				3250
42_2	ATCCAGGTCA	AGGAGGTCAC	GACGAACGAC	GGCGTTACGA	CCATCGCTAA
42_8	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
42_15	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
42_5b	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
42_1b	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
42_13	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
42_3a	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
42_4	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
42_5a	ATCCAGGTCA	AGGAGGTCAC	GACGAACGAC	GGCGTTACGA	CCATCGCTAA
42_10	ATCCAGGTCA	AGGAGGTCAC	GACGAACGAC	GGCGTTACGA	CCATCGCCAA
42_3b	ATCCAGGTCA	AGGAGGTCAC	GACGAACGAC	GGCGTTACGA	CCATCGCTAA
42_11	ATCCAGGTCA	AGGAGGTCAC	GACGAACGAC	GGCGTTACGA	CCATCGCTAA
42_6b	ATCCAGGTCA	AGGAGGTCAC	GACGGACGAC	GGCGTTACGA	CCATCGCTAA
43_1	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
43_5	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
43_12	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
43_20	ATCCAGGTCA	AGGAAGTCAC	GACGAACGAA	GGCACCAAGA	CCATCGCCAA
43_21	ATCCAGGTCA	AGGAAGTCAC	GACGAACGAA	GGCACCAAGA	CCATCGCCAA
43_23	ATCCAGGTCA	AGGAAGTCAC	GACGAACGAA	GGCACCAAGA	CCATCGCCAA
43_25	ATCCAGGTCA	AGGAAGTCAC	GACGAACGAA	GGCACCAAGA	CCATCGCCAA
44_1	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
44_5	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
223_10	ATCCAGGTCA	AGGAGGTCAC	GACGAATGAC	GGTGTACAA	CCATCGCTAA
223_2	ATCCAGGTCA	AGGAGGTCAC	GACGAATGAC	GGTGTACAA	CCATCGCTAA
223_4	ATCCAGGTCA	AGGAGGTCAC	GACGAATGAC	GGCGTCACAA	CCATCGCTAA
223_5	ATCCAGGTCA	AGGAGGTCAC	GACGAATGAC	GGCGTCACAA	CCATCGCTAA
223_6	ATCCAGGTCA	AGGAGGTCAC	GACGAATGAC	GGTGTACAA	CCATCGCTAA
223_7	ATCCAGGTCA	AGGAGGTCAC	GACGAATGAC	GGCGTCACAA	CCATCGCTAA
A3_4	ATCCAAGTCA	AGGAGGTCAC	GCAGAATGAT	GGAACCACGA	CCATCGCCAA
A3_5	ATCCAAGTCA	AGGAGGTCAC	GCAGAATGAT	GGAACCACGA	CCATCGCCAA
A3_7	ATCCAAGTCA	AGGAGGTCAC	GCAGAATGAT	GGAACCACGA	CCATCGCCAA
A3_3	ATCCAAGTCA	AGGAGGTCAC	GCAGAATGAT	GGAACCACGA	CCATCGCCAA
42_12	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
AAV1	ATCCAAGTCA	AGGAGGTCAC	GACGAATGAT	GGCGTCACAA	CCATCGCTAA
AAV2	ATTCAAGTCA	AAGAGGTCAC	GCAGAATGAC	GGTACGACGA	CGATTGCCAA
AAV3	ATCCAAGTTA	GAGGGGTCAC	GCAGAACGAT	GGCACGACGA	CTATTGCCAA
AAV8	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA
AAV9	ATCCAGGTCA	AGGAGGTTAC	GACGAACGAA	GGCACCAAGA	CCATCGCCAA
AAV7	ATCCAGGTCA	AGGAGGTCAC	GACGAATGAC	GGCGTTACGA	CCATCGCTAA
44_2	ATCCAGGTCA	AGGAGGTCAC	GCAGAATGAA	GGCACCAAGA	CCATCGCCAA

Fig. 1AAN

	3251				3300
42_2	TAACCTTACC	AGCACGATTC	AGGTCTTCTC	GGACTCGGAG	TACCAACTGC
42_8	TAACCTTACC	AGCACGATTC	AGGTCTTTAC	GGACTCGGAA	TACCAGCTCC
42_15	TAACCTTACC	AGCACGATTC	AGGTCTTTAC	GGACTCGGAA	TACCAGCTCC
42_5b	TAACCTTACC	AGCACGATTC	AGGTCTTTAC	GGACTCGGAA	TACCAGCTCC
42_1b	TAACCTTACC	AGCACGATTC	AGGTCTTTAC	GGACTCGGAA	TACCAGCTCC
42_13	TAACCTTACC	AGCACGATTC	AGGTCTTTAC	GGACTCGGAA	TACCAGCTCC
42_3a	TAACCTTACC	AGCACGATTC	AGGTCTTTAC	GGACTCGGAA	TACCAGCTCC
42_4	TAACCTTACC	AGCACGATTC	AGGTCTTTAC	GGACTCGGAA	TACCGGCTCC
42_5a	TAACCTTACC	AGCACGATTC	AGGTCTTCTC	GGACTCGGAG	TACCAACTGC
42_10	TAACCTTACC	AGCACGATTC	AGGTCTTCTC	GGACTCGGAG	TACCAACTGC
42_3b	TAACCTTACC	AGCACGATTC	AGGTCTTCTC	GGACTCGGAG	TACCAACTGC
42_11	TAACCTTACC	AGCACGATTC	AGGTCTTCTC	GGACTCGGAG	TACCAACTGC
42_6b	TAACCTTACC	AGCACGATTC	AGGTCTTCTC	GGACTCGGAG	TACCAACTGC
43_1	TAACCTTACC	AGCACGATTC	AGGTGTTTAC	GGACTCGGAA	TACCAGCTCC
43_5	TAACCTTACC	AGCACGATTC	AGGTGTTTAC	GGACTCGGAA	TACCAGCTCC
43_12	TAACCTTACC	AGCACGATTC	AGGTGTTTAC	GGACTCGGAA	TACCAGCTCC
43_20	TAATCTCACC	AGCACCGTGC	AGGTCTTTAC	GGACTCGGAG	TACCAGTTAC
43_21	TAATCTCACC	AGCACCGTGC	GGGTCTTTAC	GGACTCGGAG	TACCAGTTAC
43_23	TAATCTCACC	AGCACCGTGC	AGGTCTTTAC	GGACTTGGAG	TACCAGTTAC
43_25	TAATCTCACC	AGCACCGTGC	AGGTCTTTAC	GGACTCGGAG	TACCAGTTAC
44_1	TAACCTTACC	AGCACGATTC	AGGTCTTTAC	GGACTCGGAA	TACCAGCTCC
44_5	TAACCTTACC	AGCACGATTC	AGGTCTTTAC	GGACTCGGAA	TACCAGCTCC
223_10	TAACCTTACC	AGCACGGTTC	AGGTCTTTTC	GGACTCGGAA	TATCAACTGC
223_2	TAACCTTACC	AGCACGGTTC	AGGTCTTTTC	GGACTCGGAA	TATCAACTGC
223_4	TAACCTTACC	AGCACGGTTC	AGGTCTTTTC	GGACTCGGAA	TATCAACTGC
223_5	TAACCTTACC	AGCACGGTTC	AGGTCTTTTC	GGACTCGGAA	TATCAACTGC
223_6	TAACCTTACC	AGCACGGTTC	AGGTCTTTTC	GGACTCGGAA	TATCAACTGC
223_7	TAACCTTACC	AGCACGGTTC	AGGTCTTTTC	GGACCCGGAA	TATCAACTGC
A3_4	TAACCTTACC	AGCACGGTGC	AGGTCTTCAC	AGACTCTGAG	TACCAGCTGC
A3_5	TAACCTTACC	AGCACGGTGC	AGGTCTTCAC	AGACTCTGAG	TACCAGCTGC
A3_7	TAACCTTACC	AGCACGGTGC	AGGTCTTCAC	AGACTCTGAG	TACCAGCTGC
A3_3	TAACCTTACC	AGCGCGGTGC	AGGTCTTCAC	AGACTCTGAG	TACCAGCTGC
42_12	TAACCTTACC	AGCACGATTC	AGGTCTTTAC	GGACTCGGAA	TACCAGCTCC
AAV1	TAACCTTACC	AGCACGGTTC	AAGTCTTCTC	GGACTCGGAG	TACCAGCTTC
AAV2	TAACCTTACC	AGCACGGTTC	AGGTGTTTAC	TGACTCGGAG	TACCAGCTCC
AAV3	TAACCTTACC	AGCACGGTTC	AAGTGTTTAC	GGACTCGGAG	TATCAGCTCC
AAV8	TAACCTCACC	AGCACCATCC	AGGTGTTTAC	GGACTCGGAG	TACCAGCTGC
AAV9	TAACCTTACC	AGCACCGTCC	AGGTCTTTAC	GGACTCGGAG	TACCAGCTAC
AAV7	TAACCTTACC	AGCACGATTC	AGGTATTCTC	GGACTCGGAA	TACCAGCTGC
44_2	TAACCTTACC	AGCACGATTC	AGGTCTTTAC	GGACTCGGAA	TACCAGCTCC

Fig. 1AAO

	3301				3350
42_2	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTCCCTCC	GTTCCCTGCG
42_8	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCG
42_15	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCCGCCTCC	GTTCCCGGCG
42_5b	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCG
42_1b	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCG
42_13	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCG
42_3a	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCG
42_4	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCG
42_5a	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTCCCTCC	GTTCCCTGCG
42_10	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTCCCTCC	GTTCCCTGCG
42_3b	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTCCCTCC	GTTCCCTGCG
42_1	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTCCCTCC	GTTCCCTGCG
42_6b	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTCCCTCC	GTTCCCTGCG
43_1	CGTACGTCCC	CGGCTCTGCG	CACCAGGGCT	GCCTCCCTCC	GTTCCCGGCG
43_5	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTCCCTCC	GTTCCCGGCG
43_12	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTCCCTCC	GTTCCCGGCG
43_20	CGTACGTGCT	AGGATCCGCT	CACCAGGGAT	GTCTGCCTCC	GTTCCCGGCG
43_21	CGTACGTGCT	AGGATCCGCT	CACCAGGGAT	GTCTGCCTCC	GTTCCCGGCG
43_23	CGTACGTGCT	AGGATCCGCT	CACCAGGGAT	GTCTGCCTCC	GTTCCCGGCG
43_25	CGTACGTGCT	AGGATCCGCT	CACCAGGGAT	GTCTGCCTCC	GTTCCCGGCG
44_1	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCG
44_5	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCG
223_10	CGTACGTCCT	CGGCTCCGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCA
223_2	CGTACGTCCT	CGGCTCCGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCA
223_4	CGTACGTCCT	CGGCTCCGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCA
223_5	CGTACGTCCT	CGGCTCCGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCA
223_6	CGTACGTCCT	CGGCTCCGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCA
223_7	CGTACGTCCT	CGGCTCCGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCA
A3_4	CCTACGTCCT	CGGTTCCGGCT	CACCAGGGCT	GCCTTCCGCC	GTTCCCAGCA
A3_5	CCTACGTCCT	CGGTTCCGGCT	CACCAGGGCT	GCCTTCCGCC	GTTCCCAGCA
A3_7	CCTACGTCCT	CGGTTCCGGCT	CACCAGGGCT	GCCTTCCGCC	GTTCCCAGCA
A3_3	CCTACGTCCT	CGGTTCCGGCT	CACCAGGGCT	GCCTTCCGCC	GTTCCCAGCA
42_12	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCG
AAV1	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTCCCTCC	GTTCCCGGCG
AAV2	CGTACGTCCT	CGGCTCGGCG	CATCAAGGAT	GCCTCCCGCC	GTTCCCAGCA
AAV3	CGTACGTGCT	CGGGTCGGCG	CACCAAGGCT	GTCTCCCGCC	GTTTCCAGCG
AAV8	CGTACGTTCT	CGGCTCTGCC	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCG
AAV9	CGTACGTCCT	AGGCTCTGCC	CACCAAGGAT	GCCTGCCACC	GTTTCCTGCA
AAV7	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCG
44_2	CGTACGTCCT	CGGCTCTGCG	CACCAGGGCT	GCCTGCCTCC	GTTCCCGGCG

Fig. 1AAP

	3351				3400
42_2	GACGTGTTCA	TGATTCCTCA	GTACGGATAT	CTGACTCTAA	ACAACGGCAG
42_8	GACGTCTTCA	TGATTCCTCA	GTACGGGTAC	CTGACTCTGA	ACAACGGCAG
42_15	GACGTCTTCA	TGATTCCTCA	GTACGGGTAC	CTGACTCTGA	ACAACGGCAG
42_5b	GACGTCTTCA	TGATTCCTCA	GTACGGGTAC	CTGACTCTGA	ACAACGGCAG
42_1b	GACGTCTTCA	TGATTCCTCA	GTACGGGTAC	CTGACTCTGA	ACAACGGCAG
42_13	GACGTCTTCA	TGATTCCTCA	GTACGGGTAC	CTGACTCTGA	ACAACGGCAG
42_3a	GACGTCTTCA	TGATTCCTCA	GTACGGGTAC	CTGACTCTGA	ACAACGGCAG
42_4	GACGTCTTCA	TGATTCCTCA	GTACGGGTAC	CTGACTCTGA	ACAACGGCAG
42_5a	GACGTGTTCA	TGATTCCTCA	GTACGGATAT	CTGACTCTAA	ACAACGGCAG
42_10	GACGTGTTCA	TGATTCCTCA	GTACGGATAT	CTGACTCTAA	ACAACGGCAG
42_3b	GACGTGTTCA	TGATTCCTCA	GTACGGATAT	CTGACTCTAA	ACAACGGCAG
42_1	GACGTGTTCA	TGATTCCTCA	GTACGGATAT	CTGACTCTAA	ACAACGGCAG
42_6b	GACGTGTTCA	TGATTCCTCA	GTACGGATAT	CTGACTCTAA	ACAACGGCAG
43_1	GACGTCTTCA	TGATTCCTCA	GTACGGGTAT	CTGACCCTAA	ACAATGGCAG
43_5	GACGTCTTCA	TGATTCCTCA	GTACGGGTAT	CTGACCCTAA	ACAATGGCAG
43_12	GACGTCTTCA	TGATTCCTCA	GTACGGGTAT	CTGACCCTAA	ACAATGGCAG
43_20	GACGTCTTCA	CGGTTCCCTCA	GTACGGCTAT	TTAACTTTAA	ACAATGGAAG
43_21	GACGTCTTCA	TGGTTCCTCA	GTACGGCTAT	TTAACTTTAA	ACAATGGAAG
43_23	GACGTCTTCA	TGGTTCCTCA	GTACGGCTAT	TTAACTTTAA	ACAATGGAAG
43_25	GACGTCTTCA	TGGTTCCTCA	GTACGGCTAT	TTAACTTTAA	ACAATGGAAG
44_1	GACGTCTTCA	TGATTCCTCA	GTACGGGTAC	CTGACTCTGA	ACAATGGCAG
44_5	GACGTCTTCA	TGATTCCTCA	GTACGGGTAC	CTGACTCTGA	ACAATGGCAG
223_10	GACGTGTTCA	TGATTCCGCA	GTACGGATAC	CTGACTCTGA	ACAATGGCAG
223_2	GACGTGTTCA	TGATTCCGCA	GTACGGATAC	CTGACTCTGA	ACAATGGCAG
223_4	GACGTGTTCA	TGATTCCGCA	GTACGGATAC	CTGACTCTGA	ACAATGGCAG
223_5	GACGTGTTCA	TGATTCCGCA	GTACGGATAC	CTGACTCTGA	ACAATGGCAG
223_6	GACGTGTTCA	TGATTCCGCA	GTACGGATAC	CTGACTCTGA	ACAATGGCAG
223_7	GACGTGTTCA	TGATTCCGCA	GTACGGATAC	CTGACTCTGA	ACAATGGCAG
A3_4	GACGTCTTCA	TGATTCCTCA	GTACGGCTAC	TTGACTCTGA	ACAATGGCAG
A3_5	GACGTCTTCA	TGATTCCTCA	GTACGGCTAC	TTGACTCTGA	ACAATGGCAG
A3_7	GACGTCTTCA	TGATTCCTCA	GTACGGCTAC	TTGACTCTGA	ACAATGGCAG
A3_3	GACGTCTTCA	TGATTCCTCA	GTACGGCTAC	TTGACTCTGA	ACAATGGCAG
42_12	GACGTCTTCA	TGATTCCTCA	GTACGGGTAC	CTGACTCTGA	ACAACGGCAG
AAV1	GACGTGTTCA	TGATTCCGCA	ATACGGCTAC	CTGACGCTCA	ACAATGGCAG
AAV2	GACGTCTTCA	TGGTGCCACA	GTATGGATAC	CTCACCCTGA	ACAACGGGAG
AAV3	GACGTCTTCA	TGGTCCCTCA	GTATGGATAC	CTCACCCTGA	ACAACGGGAG
AAV8	GACGTGTTCA	TGATTCCCCA	GTACGGCTAC	CTAACACTCA	ACAACGGTAG
AAV9	GACGTCTTCA	TGGTTCCTCA	GTACGGCTAC	CTGACGCTCA	ACAATGGAAG
AAV7	GACGTCTTCA	TGATTCCTCA	GTACGGCTAC	CTGACTCTCA	ACAATGGCAG
44_2	GACGTCTTCA	TGATTCCTCA	GTACGGGTAC	CTGACTCTGA	ACAATGGCAG

Fig. 1AAQ

	3401				3450
42_2	TCAGTCTGTG	GGACGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
42_8	TCAGGCCGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
42_15	TCAGGCCGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
42_5b	TCAGGCCGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
42_1b	TCAGGCCGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
42_13	TCAGGCCGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
42_3a	TCAGGCCGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
42_4	TCAGGCCGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
42_5a	TCAGTCTGTG	GGACGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
42_10	TCAGTCTGTG	GGACGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
42_3b	TCAGTCTGTG	GGACGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
42_11	TCAGTCTGTG	GGACGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
42_6b	TCAGTCTGTG	GGACGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
43_1	TCAGGCTGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAATAC	TTCCCTTCTC
43_5	TCAGGCTGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAATAC	TTCCCTTCTC
43_12	TCAGGCTGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAATAC	TTCCCTTCTC
43_20	CCAAGCCCTG	GGACGTTCCCT	CCTTCTACTG	TCTGGAGTAT	TTCCCATCGC
43_21	CCAAGCCCTG	GGACGTTCCCT	CCTTCTACTG	TCTGGAGTAT	TTCCCATCGC
43_23	CCAAGCCCTG	GGACGTTCCCT	CCTTCTACTG	TCTGGAGTAT	TTCCCATCGC
43_25	CCAAGCCCTG	GGACGTTCCCT	CCTTCTACTG	TCTGGAGTAT	TTCCCATCGC
44_1	TCAGGCCGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
44_5	TCAGGCCGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
223_10	CCAATCGGTA	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
223_2	CCAATCGGTA	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
223_4	CCAATCGGTA	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
223_5	CCAATCGGTA	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
223_6	CCAATCGGTA	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
223_7	CCAATCGGTA	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
A3_4	CCAAGCGGTA	GGACGTTCTT	CATTCTACTG	TCTAGAGTAT	TTTCCCTCTC
A3_5	CCAAGCGGTA	GGACGTTCTT	CATTCTACTG	TCTAGAGTAT	TTTCCCTCTC
A3_7	CCAAGCGGTA	GGACGTTCTT	CATTCTACTG	TCTAGAGTAT	TTTCCCTCTC
A3_3	CCAAGCGGTA	GGACGTTCTT	CATTCTACTG	TCTAGAGTAT	TTTCCCTCTC
42_12	TCAGGCCGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC
AAV1	CCAAGCCGTG	GGACGTTTCAT	CCTTTTACTG	CCTGGAATAT	TTCCCTTCTC
AAV2	TCAGGCAGTA	GGACGCTCCT	CATTTTACTG	CCTGGAGTAC	TTTCCTTCTC
AAV3	TCAAGCGGTG	GGACGCTCAT	CCTTTTACTG	CCTGGAGTAC	TTTCCCTTCGC
AAV8	TCAGGCCGTG	GGACGCTCCT	CCTTCTACTG	CCTGGAATAC	TTTCCTTCTC
AAV9	TCAAGCGTTA	GGACGTTCTT	CCTTCTACTG	TCTGGAATAC	TTCCCTTCTC
AAV7	TCAGTCTGTG	GGACGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTCCCTTCTC
44_2	TCAGGCCGTG	GGCCGTTCCCT	CCTTCTACTG	CCTGGAGTAC	TTTCCTTCTC

Fig. 1AAR

	3451				3500
42_2	AGATGCTGAG	AACGGGCAAT	AACTTTGAAT	TCAGCTACAC	CTTTGAGGAA
42_8	AAATGCTGAG	AACGGGCAAC	AACTTTGAGT	TCAGCTACCA	GTTTGAGGAC
42_15	AAATGCTGAG	AACGGGCAAC	AACTTTGAGT	TCAGCTACCA	GTTTGAGGAC
42_5b	AAATGCTGAG	AACGGGCAAC	AACTTTGAGT	TCAGCTACCA	GTTTGAGGAC
42_1b	AAATGCTGAG	AACGGGCAAC	AACTTTGAGT	TCAGCTACCA	GTTTGAGGAC
42_13	AAATGCTGAG	AACGGGCAAC	AACTTTGAGT	TCAGCTACCA	GTTTGAGGAC
42_3a	AAATGCTGAG	AACGGGCAAC	AACTTTGAGT	TCAGCTACCA	GTTTGAGGAC
42_4	AAATGCTGAG	AACGGGCAAC	AACTTTGAGT	TCAGCTACCA	GTTTGAGGAC
42_5a	AGATGCTGAG	AACGGGCAAT	AACTTTGAAT	TCAGCTACCA	GTTTGAGGAC
42_10	AGATGCTGAG	AACGGGCAAT	AACTTTGAAT	TCAGCTACAC	CTTTGAGGAA
42_3b	AGATGCTGAG	AACGGGCAAT	AACTTTGAAT	TCAGCTACAC	CTTTGAGGAA
42_11	AGATGCTGAG	AACGGGCAAT	AACTTTGAAT	TCAGCTACAC	CTTTGAGGAA
42_6b	AGATGCTGAG	AACGGGCAAT	AACTTTGAAT	TCAGCTACAC	CTTTGAGGAA
43_1	AAATGCTGAG	GACGGGCAAC	AACTTTGAAT	TCAGCTACAC	CTTCGAGGAC
43_5	AAATGCTGAG	GACGGGCAAC	AACTTTGAAT	TCAGCTACAC	CTTCGAGGAC
43_12	AAATGCTGAG	GACGGGCAAC	AACTTTGAAT	TCAGCTACAC	CTTCGAGGAC
43_20	AGATGCTGAG	AACCGGCAAC	AACTTTCAGT	TCAGCTACAC	CTTCGAGGAC
43_21	AGATGCTGAG	AACCGGCAAC	AACTTTCAGT	TCAGCTACAC	CTTCGAGGAC
43_23	AGATGCTGAG	AACCGGCAAC	AACTTTCAGT	TCAGCTACAC	CTTCGAGGAC
43_25	AGATGCTGAG	AACCGGCAAC	AACTTTCAGT	TCAGCTACAC	CTTCGAGGAC
44_1	AAATGCTGAG	AACGGGCAAC	AACTTTGAGT	TCAGCTACCA	GTTTGAGGAC
44_5	AAATGCTGAG	AACGGGCAAC	AACTTTGAGT	TCAGCTACCA	GTTTGAGGAC
223_10	AGATGCTGAG	AACGGGCAAC	AACTTCACCT	TTAGCTACAC	CTTCGAGGAC
223_2	AGATGCTGAG	AACGGGCAAC	AACTTCACCT	TTAGCTACAC	CTTCGAGGAC
223_4	AGATGCTGAG	AACGGGCAAC	AACTTCACCT	TTAGCTACAC	CTTCGAGGAC
223_5	AGATGCTGAG	AACGGGCAAC	AACTTCACCT	TTAGCTACAC	CTTCGAGGAC
223_6	AGATGCTGAG	AACGGGCAAC	AACTTCACCT	TTAGCTACAC	CTTCGAGGAC
223_7	AGATGCTGAG	AACGGGCAAC	AACTTCACCT	TTAGCTACAC	CTTCGAGGAC
A3_4	AGATGCTGAG	GACGGGAAAC	AACTTCACCT	TCAGCTACAC	TTTTGAAGAC
A3_5	AGATGCTGAG	GACGGGAAAC	AACTTCACCT	TCAGCTACAC	TTTTGAAGAC
A3_7	AGATGCTGAG	GACGGGAAAC	AACTTCACCT	TCAGCTACAC	TTTTGAAGAC
A3_3	AGATGCTGAG	GACGGGAAAC	AACTTCACCT	TCAGCTACAC	TTTTGAAGAC
42_12	AAATGCTGAG	AACGGGCAAC	AACTTTGAGT	TCAGCTACCA	GTTTGAGGAC
AAV1	AGATGCTGAG	AACGGGCAAC	AACTTTACCT	TCAGCTACAC	CTTTGAGGAA
AAV2	AGATGCTGCG	TACCGGAAAC	AACTTTACCT	TCAGCTACAC	TTTTGAGGAC
AAV3	AGATGCTAAG	GACTGGAAAT	AACTTCCAAT	TCAGCTATAC	CTTCGAGGAT
AAV8	AGATGCTGAG	AACCGGCAAC	AACTTCCAGT	TTACTTACAC	CTTCGAGGAC
AAV9	AGATGCTGAG	AACCGGCAAC	AACTTTCAGT	TCAGCTACAC	TTTCGAGGAC
AAV7	AGATGCTGAG	AACGGGCAAC	AACTTTGAGT	TCAGCTACAG	CTTCGAGGAC
44_2	AAATGCTGAG	AACGGGCAAC	AACTTTGAGT	TCAGCTACCA	GTTTGAGGAC

Fig. 1AAS

	3501				3550
42_2	GTGCCTTTCC	ACAGCAGCTA	TGCGCACAGC	CAGAGCCTGG	ACCGGCTGAT
42_8	GTGCCTTTTC	ACAGCAGCTA	CGCGCACAGC	CAAAGCCTGG	ACCGGCTGAT
42_15	GTGCCTTTTC	ACAGCAGCTA	CGCGCATAGC	CAAAGCCTGG	ACCGGCTGAT
42_5b	GTGCCTTTTC	ACAGCAGCTA	CGCGCACAGC	CAAAGCCTGG	ACCGGCTGAT
42_1b	GTGCCTTTTC	ACAGCAGCTA	TGCGCACAGC	CAAAGCCTGG	ACCGGCTGAT
42_13	GTGCCTTTTC	ACAGCAGCTA	TGCGCACAGC	CAAAGCCTGG	ACCGGCTGAT
42_3a	GTGCCTTTTC	ACAGCAGCTA	CGCGCACAGC	CAAAGCCTGG	ACCGGCTGAT
42_4	GTGCCTTTTC	ACAGCAGCTA	CGCGCACAGC	CAAAGCCTGG	ACCGGCTGAT
42_5a	GTGCCCTTTC	ACAGCAGCTA	CGCGCACAGC	CAAAGCCTGG	ACCGGCTGAT
42_10	GTGCCTTTCC	ACAGCAGCTA	TGCGCACAGC	CAGAGCCTGG	ACCGGCTGAT
42_3b	GTGCCTTTCC	ACAGCAGCTA	TGCGCACAGC	CAGAGCCTGG	ACCGGCTGAT
42_11	GTGCCTTTCC	ACAGCAGCTA	TGCGCACAGC	CAGAGCCTGG	ACCGGCTGAT
42_6b	GTGCCTTTCC	ACAGCAGCTA	TGCGCATAGC	CAGAGCCTGG	ACCGGCTGAT
43_1	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGCCTGG	ACCGGCTGAT
43_5	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGCCTGG	ACCGGCTGAT
43_12	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGCCTGG	ACCGGCTGAT
43_20	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGCCTGG	ACAGGCTGAT
43_21	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGCCTGG	ACAGGCTGAT
43_23	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGCCTGG	ACAGGCTGAT
43_25	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGCCTGG	ACAGGCTGAT
44_1	GTGCCTTTTC	ACAGCAGCTA	CGCGCACAGC	CAAAGCCTGG	ACCGGCTGAT
44_5	GTGCCTTTTC	ACAGCAGCTA	CGCGCACAGC	CAAAGCCTGG	ACCGGCTGAT
223_10	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGTCTGG	ACCGGCTGAT
223_2	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGTCTGG	ACCGGCTGAT
223_4	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGTCTGG	GCCGGCTGAT
223_5	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGTCTGG	GCCGGCTGAT
223_6	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGTCTGG	ACCGGCTGAT
223_7	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGTCTGG	ACCGGCTGAT
A3_4	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGTCTGG	ATCGGCTGAT
A3_5	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGTCTGG	ATCGGCTGAT
A3_7	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGTCTGG	ATCGGCTGAT
A3_3	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGTCTGG	ATCGGCTGAT
42_12	GTGCCTTTTC	ACAGCAGCTA	CGCGCACAGC	CAAAGCCTGG	ACCGGCTGAC
AAV1	GTGCCTTTCC	ACAGCAGCTA	CGCGCACAGC	CAGAGCCTGG	ACCGGCTGAT
AAV2	GTTCCTTTCC	ACAGCAGCTA	CGCTCACAGC	CAGAGTCTGG	ACCGTCTCAT
AAV3	GTACCTTTTC	ACAGCAGCTA	CGCTCACAGC	CAGAGTTTGG	ATCGCTTGAT
AAV8	GTGCCTTTCC	ACAGCAGCTA	CGCCACAGC	CAGAGCTTGG	ACCGGCTGAT
AAV9	GTGCCTTTCC	ACAGCAGCTA	CGCACACAGC	CAGAGTCTAG	ATCGACTGAT
AAV7	GTGCCTTTCC	ACAGCAGCTA	CGCACACAGC	CAGAGCCTGG	ACCGGCTGAT
44_2	GTGCCTTTTC	ACAGCAGCTA	CGCGCACAGC	CAAAGCCTGG	ACCGGCTGAT

Fig. 1AAT

	3551				3600
42_2	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGGCCCCG	ACCCAGAGCA
42_8	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGTCTCGG	ACTCAGTCCA
42_15	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGTCTCGG	ACTCAGTCCA
42_5b	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGTCTCGG	ACTCAGTCCA
42_1b	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGTCTCGG	ACTCAGTCCA
42_13	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGTCTCGG	ACTCAGTCCA
42_3a	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGTCTCGG	ACTCAGTCCA
42_4	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGTCTCGG	ACTCAGTCCA
42_5a	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGTCTCGG	ACTCAGTCCA
42_10	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGGCCCCG	ACCCAGAGCA
42_3b	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGGCCCCG	ACCCAGAGCA
42_11	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGGCCCCG	ACCCAGAGCA
42_6b	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGGCCCCG	ACCCAGAGCA
43_1	GAACCCCTC	ATCGACCAGT	ACCTGTATTA	CTTATCCAGA	ACTCAGTCCA
43_5	GAACCCCTC	ATCGACCAGT	ACCTGTATTA	CTTATCCAGA	ACTCAGTCCA
43_12	GAACCCCTC	ATCGACCAGT	ACCTGTATTA	CTTATCCAGA	ACTCAGTCCA
43_20	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGGTCAGA	ACGCAAACGA
43_21	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGGTCAGA	ACGCAAACGA
43_23	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGGTCAGA	ACGCAAACGA
43_25	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGGTCAGA	ACGCAAACGA
44_1	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGTCTCGG	ACTCAGTCCA
44_5	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGTCTCGG	ACTCAGTCCA
223_10	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CTTGGCCAGA	ACACAGAGCA
223_2	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CTTGGCCAGA	ACACAGAGCA
223_4	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CTTGGCCAGA	ACACAGAGCA
223_5	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CTTGGCCAGA	ACACAGAGCA
223_6	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CTTGGCCAGA	ACACAGAGCA
223_7	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CTTGGCCAGA	ACACAGAGCA
A3_4	GAATCCTCTC	ATTGACCAGT	ACCTGTATTA	CCTGAGCAAA	ACTCAGGGTA
A3_5	GAATCCTCTC	ATTGACCAGT	ACCTGTATTA	CCTGAGCAAA	ACTCAGGGTA
A3_7	GAATCCTCTC	ATTGACCAGT	ACCTGTATTA	CCTGAGCAAA	ACTCAGGGTA
A3_3	GAATCCTCTC	ATTGACCAGT	ACCTGTATTA	CCTGAGCAAA	ACTCAGGGTA
42_12	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGGCCCCG	ACCCAGAGCA
AAV1	GAATCCTCTC	ATCGACCAAT	ACCTGTATTA	CCTGAACAGA	ACTCAA.AT
AAV2	GAATCCTCTC	ATCGACCAGT	ACCTGTATTA	CTTGAGCAGA	ACAAACACTC
AAV3	GAATCCTCTT	ATTGATCAGT	ATCTGTACTA	CCTGAACAGA	ACGCAAGGAA
AAV8	GAATCCTCTG	ATTGACCAGT	ACCTGTACTA	CTTGTCTCGG	ACTCAAACAA
AAV9	GAATCCCCTC	ATCGACCAGT	ACCTATACTA	CCTGGTCAGA	ACACAGACAA
AAV7	GAATCCCCTC	ATCGACCAGT	ACTTGTACTA	CCTGGCCAGA	ACACAGAGTA
44_2	GAATCCCCTC	ATCGACCAGT	ACCTGTACTA	CCTGTCTCGG	ACTCAGTCCA

Fig. 1AAU

	3601				3650
42_2	CTACGG..GG	TCCACAAGGG	AGCTGCA.GT	TCCA.....	TCAGGCTGGG
42_8	CGGGA...GG	TACCGCAGGA	ACTCAGCAGT	TGCTATTTTC	TCAGGCCGGG
42_15	CGGGA...GG	TACCGCAGGA	ACTCAGCAGT	TGCTATTTTC	TCAGGCCGGG
42_5b	CGGGA...GG	TACCGCAGGA	ACTCAGCAGT	TGCTATTTTC	TCAGGCCGGG
42_1b	CGGGA...GG	TACCGCAGGA	ACTCAGCAGT	TGCTATTTTC	TCAGGCCGGG
42_13	CGGGA...GG	TACCGCAGGA	ACTCAGCAGT	TGCTATTTTC	TCAGGCCGGG
42_3a	CGGGA...GG	TACCGCAGGA	ACTCAGCAGT	TGCTATTTTC	TCAGGCCGGG
42_4	CGGGA...GG	TACCGCAGGA	ACTCAGCAGT	TGCTATTTTC	TCAGGCCGGG
42_5a	CGGGA...GG	TACCGCAGGA	ACTCAGCAGT	TGCTATTTTC	TCAGGCCGGG
42_10	CTACG...GG	GTCCACAAGG	GAGCTGCAGT	TCCA.....	TCAGGCTGGG
42_3b	CTACG...GG	GTCCACAAGG	GAGCTGCAGT	TCCA.....	TCAGGCTGGG
42_11	CTACG...GG	GTCCACAAGG	GAGCTGCAGT	TCCA.....	TCAGGCTGGG
42_6b	CTACG...GG	GTCCACAAGG	GAGCTGCAGT	TCCA.....	TCAGGCTGGG
43_1	CAGGA...GG	AACTCAAGGT	ACTCAGCAAT	TGTTATTTTC	TCAAGCCGGG
43_5	CAGGA...GG	AACTCAAGGT	ACTCAGCAAT	TGTTATTTTC	TCAAGCCGGG
43_12	CAGGA...GG	AACTCAAGGT	ACTCAGCAAT	TGTTATTTTC	TCAAGCCGGG
43_20	CT.....GG	AACTGGAGGG	ACGCAGACTC	TGGCATTTCAG	CCAAGCCGGT
43_21	CT.....GG	AACTGGAGGG	ACGCAGACTC	TGGCATTTCAG	CCAAGCCGGT
43_23	CT.....GG	AACTGGAGGG	ACGCAGACTC	TGGCATTTCAG	CCAAGCCGGT
43_25	CT.....GG	AACTGGAGGG	ACGCAGACTC	TGGCATTTCAG	CCAAGCCGGT
44_1	CGGGA...GG	TACCGCAGGA	ACTCAGCAGT	TGCTATTTTC	TCAGGCCGGG
44_5	CGGGA...GG	TACCGCAGGA	ACTCAGCAGT	TGCTATTTTC	TCAGGCCGGG
223_10	ACGCAGGAGG	TACTGCTGGC	AATCGGGAAC	TGCAGTTTTA	TCAGGGCGGA
223_2	ACGCAGGAGG	TACTGCTGGC	AATCGGGAAC	TGCAGTTTTA	TCAGGGCGGA
223_4	ACGCAGGAGG	TACTGCTGGC	AATCGGGAAC	TGCAGTTTTA	TCAGGGCGGA
223_5	ACGCAGGAGG	TACTGCTGGC	AATCGGGAAC	TGCAGTTTTA	TCAGGGCGGA
223_6	ACGCAGGAGG	TACTGCTGGC	AATCGGGAAC	TGCAGTTTTA	TCAGGGCGGA
223_7	ACGCAGGAGG	TACTGCTGGC	AATCGGGAAC	TGCAGTTTTA	TCAGGGCGGA
A3_4	CAAG...TGG	AACAACGCAG	CAATCGAGAC	TGCAGTTCAG	CCAAGCTGGG
A3_5	CAAG...TGG	AACAACGCAG	CAATCGAGAC	TGCAGTTCAG	CCAAGCTGGG
A3_7	CAAG...TGG	AACAACGCAG	CAATCGAGAC	TGCAGTTCAG	CCAAGCTGGG
A3_3	CAAG...TGG	AACAACGCAG	CAATCGAGAC	TGCAGTTCAG	CCAAGCTGGG
42_12	CTACG...GG	GTCCACAAGG	GGGCTGCAGT	TCCA.....	TCAGGCTGGG
AAV1	CAGTCC..GG	AAGTGCCCAA	AACAAGGACT	TGCTGTTTTC	CCGTGGGTCT
AAV2	CAAG...TGG	AACCACCACG	CAGTCAAGGC	TTCAGTTTTTC	TCAGGCCGGA
AAV3	CAACCTCTGG	AACAACCAAC	CAATCACGGC	TGCTTTTTTC	CCAGGCTGGG
AAV8	CAGGAG..GC	.ACGGCAAAT	ACGCAGACTC	TGGGCTTCAG	CCAAGGTGGG
AAV9	CTGGA.....	.ACTGGGGGA	ACTCAAACCT	TGGCATTTCAG	CCAAGCAGGC
AAV7	ACCCAGGAGG	CACAGCTGGC	AATCGGGAAC	TGCAGTTTTA	CCAGGGCGGG
44_2	CGGGA...GG	TACCGCAGGA	ACTCAGCAGT	TGCTATTTTC	TCAGGCCGGG

Fig. 1AAV

	3651				3700
42_2	CCCAACACCA	TGGCCGAGCA	ATCAAAGAAC	TGGCTGCCCCG	GACCCTGTTA
42_8	CCTAATAACA	TGTCGGCTCA	GGCCAAAAAC	TGGCTACCCG	GGCCCTGCTA
42_15	CCTAATAACA	TGTCGGCTCA	GGCCAAAAAC	TGGCTACCCG	GGCCCTGCTA
42_5b	CCTAATAACA	TGTCGGCTCA	GGCCAAAAAC	TGGCTACCCG	GGCCCTGCTA
42_1b	CCTAATAACA	TGTCGGCTCA	GGCCAAAAAC	TGGCTACCCG	GGCCCTGCTA
42_13	CCTAATAACA	TGTCGGCTCA	GGCCAAAAAC	TGGCTACCCG	GGCCCTGCTA
42_3a	CCTAATAACA	TGTCGGCTCA	GGCCAAAAAC	TGGCTACCCG	GGCCCTGCTA
42_4	CCTAATAACA	TGTCGGCTCA	GGCCAAAAAC	TGGCTACCCG	GGCCCTGCTA
42_5a	CCTAATAACA	TGTCGGCTCA	GGCCAAAAAC	TGGCTACCCG	GGCCCTGCTA
42_10	CCCAACACCA	TGGCCGAGCA	ATCAAAGAAC	TGGCTGCCCCG	GACCCTGTTA
42_3b	CCCAACACCA	TGGCCGAGCA	ATCAAAGAAC	TGGCTGCCCCG	GACCCTGTTA
42_11	CCCAACACCA	TGGCCGAGCA	ATCAAAGAAC	TGGCTGCCCCG	GACCCTGTTA
42_6b	CCCAACACCA	TGGCCGAGCA	ATCAAAGAAC	TGGCTGCCCCG	GACCCTGTTA
43_1	CCCGCAAACA	TGTCGGCTCA	GGCCAAGAAC	TGGCTACCTG	GACCGTGTTA
43_5	CCCGCAAACA	TGTCGGCTCA	GGCCAAGAAC	TGGCTACCTG	GACCGTGTTA
43_12	CCCGCAAACA	TGTCGGCTCA	GGCCAAGAAC	TGGCTACCTG	GACCGTGTTA
43_20	CCTAGCTCAA	TGGCCAACCA	GGCTAGAAAT	TGGGTGCCCCG	GACCTTGCTA
43_21	CCTAGCTCAA	TGGCCAACCA	GGCTAGAAAT	TGGGTGCCCCG	GACCTTGCTA
43_23	CCTAGCTCAA	TGGCCAACCA	GGCTAGAAAT	TGGGTGCCCCG	GACCTTGCTA
43_25	CCTAGCTCAA	TGGCCAACCA	GGCTAGAAAT	TGGGTGCCCCG	GACCTTGCTA
44_1	CCTAATAACA	TGTCGGCTCA	GGCCAAAAAC	TGGCTACCCG	GGCCCTGCTA
44_5	CCTAATAACA	TGTCGGCTCA	GGCCAAAAAC	TGGCTACCCG	GGCCCTGCTA
223_10	CCTACCACCA	TGGCCGAACA	AGCAAAGAAC	TGGCTGCCCCG	GACCTTGCTT
223_2	CCTACCACCA	TGGCCGAACA	AGCAAAGAAC	TGGCTGCCCCG	GACCTTGCTT
223_4	CCTACCACCA	TGGCCGAACA	AGCAAAGAAC	TGGCTGCCCCG	GACCTTGCTT
223_5	CCTACCACCA	TGGCCGAACA	AGCAAAGAAC	TGGCTGCCCCG	GACCTTGCTT
223_6	CCTACCACCA	TGGCCGAACA	AGCAAAGAAC	TGGCTGCCCCG	GACCTTGCTT
223_7	CCTACCACCA	TGGCCGAACA	AGCAAAGAAC	TGGCTGCCCCG	GACCTTGCTT
A3_4	CCTAGCTCCA	TGGCTCAGCA	GGCCAAAAAC	TGGCTACCGG	GACCCAGCTA
A3_5	CCTAGCTCCA	TGGCTCAGCA	GGCCAAAAAC	TGGCTACCGG	GACCCAGCTA
A3_7	CCTAGCTCCA	TGGCTCAGCA	GGCCAAAAAC	TGGCTACCGG	GACCCAGCTA
A3_3	CCTAGCTCCA	TGGCTCAGCA	GGCCAAAAAC	TGGCTACCGG	GACCCAGCTA
42_12	CCCAACACCA	TGGCCGAGCA	ATCAAAGAAC	TGGCTGCCCCG	GACCCTGTTA
AAV1	CCAGCTGGCA	TGTCTGTTCA	GGCCAAAAAC	TGGCTACCTG	GACCCTGTTA
AAV2	GCGAGTGACA	TTCTGGGACCA	GTCTAGGAAC	TGGCTTCCTG	GACCCTGTTA
AAV3	CCTCAGTCTA	TGTCTTTGCA	GGCCAGAAAT	TGGCTACCTG	GGCCCTGCTA
AAV8	CCTAATACAA	TGGCCAATCA	GGCAAAGAAC	TGGCTGCCAG	GACCCTGTTA
AAV9	CCTAGCTCAA	TGGCCAATCA	GGCTAGAAAC	TGGGTACCCG	GGCCTTGCTA
AAV7	CCTTCAACTA	TGGCCGAACA	AGCCAAGAAT	TGGTTACCTG	GACCTTGCTT
44_2	CCTAATAACA	TGTCGGCTCA	GGCCAAAAAC	TGGCTACCCG	GGCCCTGCTA

Fig. 1AAW

	3701				3750
42_2	TCGGCAGCAG	AGACTGTCAA	AAAACATAGA	CAGCAACAAC	AACAGTAACT
42_8	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
42_15	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
42_5b	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
42_1b	CCGGCAGCAA	CGCGTCTCCA	CGACAGTGTC	GCAAAATAAC	AACAGCAACT
42_13	CCGGCAGCAA	CGCGTCTCCA	CGACAGTGTC	GCAAAATAAC	AACAGCAACT
42_3a	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
42_4	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
42_5a	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
42_10	TCGGCAGCAG	AGACTGTCAA	AAAACATAGA	CAGCAACAAC	AACAGTAACT
42_3b	TCGGCAGCAG	AGACTGTCAA	AAAACATAGA	CAGCAACAAC	ACCAGTAACT
42_11	TCGGCGGCAG	AGACTGTCAA	AAGACATAGA	CAGCAACAAC	AACAGTAACT
42_6b	TCGGCAGCAG	AGACTGTCAA	AAAACATAGA	CAGCAACAAC	AACAGTAACT
43_1	CCGTCAGCAA	CGAGTTTCCA	CGACACTGTC	GCAAAACAAC	AACAGCAATT
43_5	CCGTCAGCAA	CGAGTTTCCA	CGACACTGTC	GCAAAACAAC	AACAGCAATT
43_12	CCGTCAGCAA	CGAGTTTCCA	CGACACTGTC	GCAAAACAAC	AACAGCAATT
43_20	CCGGCAGCAG	CGCGTCTCCA	CGACAACCAA	CCAGAACAAC	AACAGCAACT
43_21	CCGGCAGCAG	CGCGTCTCCA	CGACAACCAA	CCAGAGCAAC	AACAGCAACT
43_23	CCGGCAGCAG	CGCGTCTCCA	CGACAACCAA	CCAGAACAAC	AACAGCAACT
43_25	CCGGCAGCAG	CGCGTCTCCA	CGACAACCAA	CCAGAACAAC	AACAGCAACT
44_1	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
44_5	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
223_10	CCGGCAACAG	AGAGTATCCA	AGACGCTGGA	TCAAAATAAC	AACAGCAACT
223_2	CCGGCAACAG	AGAGTATCCA	AGACGCTGGA	TCAAAATAAC	AACAGCAACT
223_4	CCGGCAACAG	AGAGTATCCA	AGACGCTGGA	TCAAAATAAC	AACAGCAACT
223_5	CCGGCAACAG	AGAGTATCCA	AGACGCTGGA	TCAAAATAAC	AACAGCAACT
223_6	CCGGCAACAG	AGAGTATCCA	AGACGCTGGA	TCAAAATAAC	AACAGCAACT
223_7	CCGGCAACAG	AGAGTATCCA	AGACGCTGGA	TCAAAATAAC	AACAGCAACT
A3_4	CCGACAGCAG	CGAATGTCTA	AGACGGCTAA	TGACAACAAC	AACAGTGAAT
A3_5	CCGACAGCAG	CGAATGTCTA	AGACGGCTAA	TGACAACAAC	AACAGTGAAT
A3_7	CCGACAGCAG	CGAATGTCTA	AGACGGCTAA	TGACAACAAC	AACAGTGAAT
A3_3	CCGACAGCAG	CGAATGTCTA	AGACGGCTAA	TGACAACAAC	AACAGTGAAT
42_12	TCGGCAGCAG	AGACTGTCAA	AAAACATAGA	CAGCAACAAC	AACAGTAACT
AAV1	TCGGCAGCAG	CGCGTTTCTA	AAACAAAAAC	AGACAACAAC	AACAGCAATT
AAV2	CCGCCAGCAG	CGAGTATCAA	AGACATCTGC	GGATAACAAC	AACAGTGAAT
AAV3	CCGGCAACAG	AGACTTTCAA	AGACTGCTAA	CGACAACAAC	AACAGTAACT
AAV8	CCGCCAACAA	CGCGTCTCAA	CGACAACCGG	GCAAAACAAC	AATAGCAACT
AAV9	CCGTCAGCAG	CGCGTCTCCA	CAACCACCAA	CCAAAATAAC	AACAGCAACT
AAV7	CCGGCAACAA	AGAGTCTCCA	AAACGCTGGA	TCAAAACAAC	AACAGCAACT
44_2	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT

Fig. 1AAX

	3751				3800
42_2	TTGCCTGGAC	CGGGGCCACT	AAATACCATC	TGAATGGTAG	AAATTCATTA
42_8	TTGCTTGGAC	CGGTGCCACC	AAGTATCATC	TGAATGGCAG	AGACTCTCTG
42_15	TTGCTTGGAC	CGGTGCCACC	AAGTATCATC	TGAATGGCAG	AGACTCTCTG
42_5b	TTGCTTGGAC	CGGTGCCACC	AAGTATCATC	TGAATGGCAG	AGACTCTCTG
42_1b	TTGCTTGGAC	CGGTGCCACC	AAGTATCATC	TGAATGGCAG	AGACTCTCTG
42_13	TTGCTTGGAC	CGGTGCCACC	AAGTATCATC	TGAATGGCAG	AGACTCTCTG
42_3a	TTGCTTGGAC	CGGTGCCACC	AAGTATCATC	TGAATGGCAG	AGACTCTCTG
42_4	TTGCTTGGAC	CGGTGCCACC	AAGTATCATC	TGAATGGCAG	AGACTCTCTG
42_5a	TTGCTTGGAC	CGGTGCCACC	AAGTATCATC	TGAATGGCAG	AGACTCTCTG
42_10	TTGCCTGGAC	CGGGGCCACT	AAATACCATC	TGAATGGTAG	AAATTCATTA
42_3b	TTGCCTGGAC	CGGGGCCACT	AAATACCATC	TGAATGGTAG	AAATTCATTA
42_11	TTGCCTGGAC	CGGGGCCACT	AAATACCATC	TGAATGGTAG	AAATTCATTA
42_6b	TTGCCTGGAC	CGGGGCCACT	AAATACCATC	TGAATGGTAG	AAATTCATTA
43_1	TTGCTTGGAC	CGGTGCCACC	AAGTATCACC	TGAATGGCAG	AGACTCCCTG
43_5	TTGCTTGGAC	CGGTGCCACC	AAGTATCACC	TGAATGGCAG	AGACTCCCTG
43_12	TTGCTTGGAC	CGGTGCCACC	AAGTATCACC	TGAATGGCAG	AGACTCCCTG
43_20	TTGCCTGGAC	GGGAGCTGCC	AAGTTTAAAGC	TGAACGGCCG	AGACTCTCTA
43_21	TTGCCTGGAC	GGGAGCTGCC	AAGTTTAAAGC	TGAACGGCCG	AGACTCTCTA
43_23	TTGCCTGGAC	GGGAGCTGCC	AAGTTTAAAGC	TGAACGGCCG	AGACTCTCTA
43_25	TTGCCTGGAC	GGGAGCTGCC	AAGTTTAAAGC	TGAACGGCCG	AGACTCTCTA
44_1	TTGCCTGGAC	CGGTGCCACC	AAGTATCATC	TGAATGGCAG	AGACTCTCTG
44_5	TTGCCTGGAC	CGGTGCCACC	AAGTATCATC	TGAATGGCAG	AGACTCTCTG
223_10	TTGCCTGGAC	TGGTGCCACA	AAATACCATT	TAAATGGAAG	AAATTCATTG
223_2	TTGCCTGGAC	TGGTGCCACA	AAATACCATT	TAAATGGAAG	AAATTCATTG
223_4	TTGCCTGGAC	TGGTGCCACA	AAATACCATT	TAAATGGAAG	AAATTCATTG
223_5	TTGCCTGGAC	TGGTGCCACA	AAATACCATT	TAAATGGAAG	AAATTCATTG
223_6	TTGCCTGGAC	TGGTGCCACA	AAATACCATT	TAAATGGAAG	AAATTCATTG
223_7	TTGCCTGGAC	TGGTGCCACA	AAATACCATT	TAAATGGAAG	AAATTCATTG
A3_4	TTGCTTGGAC	TGCAGCCACC	AAATATTACC	TGAATGGAAG	AAATTCTCTG
A3_5	TTGCTTGGAC	TGCAGCCACC	AAATATTACC	TGAATGGAAG	AAATTCTCTG
A3_7	TTGCTTGGAC	TGCAGCCACC	AAATATTACC	TGAATGGAAG	AAATTCTCTG
A3_3	TTGCTTGGAC	TGCAGCCACC	AAATATTACC	TGAATGGAAG	AAATTCTCTG
42_12	TTGCCTGGAC	CGGGGCCACT	AAATACCATC	TGAATGGTAG	AAATTCATTA
AAV1	TTACCTGGAC	TGGTGCTTCA	AAATATAACC	TCAATGGGCG	TGAATCCATC
AAV2	ACTCGTGGAC	TGGAGCTACC	AAGTACCACC	TCAATGGCAG	AGACTCTCTG
AAV3	TTCCCTGGAC	AGCGGCCAGC	AAATATCATC	TCAATGGCCG	CGACTCGCTG
AAV8	TTGCCTGGAC	TGCTGGGACC	AAATACCATC	TGAATGGAAG	AAATTCATTG
AAV9	TTGCGTGGAC	GGGAGCTGCT	AAATTCAAGC	TGAACGGGAG	AGACTCGCTA
AAV7	TTGCTTGGAC	TGGTGCCACC	AAATATCACC	TGAACGGCAG	AAACTCGTTG
44_2	TTGCCTGGAC	CGGTGCCACC	AAGTATCATC	TGAATGGCAG	AGACTCTCTG

Fig. 1AAY

	3801				3850
42_2	ACCAACCCCG	GCGTAGCCAT	GGCCACCAAC	AAGGACGACG	AGGACCAGTT
42_8	GTAAATCCCG	GTGTCGCTAT	GGCAACGCAC	AAGGACGACG	AAGAGCGATT
42_15	GTAAATCCCG	GTGTCGCTAT	GGCAACGCAC	AAGGACGACG	AAGAGCGATT
42_5b	GTAAATCCCG	GTGTCGCTAT	GGCAACGCAC	AAGGACGACG	AAGAGCGATT
42_1b	GTAAATCCCG	GTGTCGCTAT	GGCAACGCAC	AAGGGCGACG	AAGAGCGATT
42_13	GTAAATCCCG	GTGTCGCTAT	GGCAACGCAC	AAGGGCGACG	AAGAGCGATT
42_3a	GTAAATCCCG	GTGTCGCTAT	GGCAACGCAC	AAGGACGACG	AAGAGCGATT
42_4	GTAAATCCCG	GTGTCGCTAT	GGCAACGCAC	AAGGACGACG	AAGAGCGATT
42_5a	GTAAATCCCG	GTGTCGCTAT	GGCAACGCAC	AAGGACGACG	AAGAGCGATT
42_10	ACCAACCCCG	GCGTAGCCAT	GGCCACCAAC	AAGGACGACG	AGGACCAGTT
42_3b	ACCAACCCCG	GCGTAGCCAT	GGCCACCAAC	AAGGACGACG	AGGACCAGTT
42_11	ACCAACCCCG	GCGTAGCCAT	GGCCACCAAC	AAGGACGACG	AGGACCAGTT
42_6b	ACCAACCCCG	GCGTAGCCAT	GGCCACCAAC	AAGGACGACG	AGGACCAGTT
43_1	GTTAATCCCG	GCGTTGCCAT	GGCTACCCAC	AAGGACGACG	AGGAGCGCTT
43_5	GTTAATCCCG	GCGTTGCCAT	GGCTACCCAC	AAGGACGACG	AGGAGCGCTT
43_12	GTTAATCCCG	GCGTTGCCAT	GGCTACCCAC	AAGGACGACG	AGGAGCGCTT
43_20	ATGAATCCGG	GCGTGGCAAT	GGCTTCCCAC	AAGGATGACG	ACGACCGCTT
43_21	ATGAATCCGG	GCGTGGCAAT	GGCTTCCCAC	AAGGATGACG	ACGACCGCTT
43_23	ATGAATCCGG	GCGTGGCAAT	GGCTTCCCAC	AAGGATGACG	ACGACCGCTT
43_25	ATGAATCCGG	GCGTGGCAAT	GGCTTCCCAC	AAGGATGACG	ACGACCGCTT
44_1	GTAAATCCCG	GTGTCGCTAT	GGCAACCCAC	AAGGACGACG	AAGAGCGATT
44_5	GTAAATCCCG	GTGTCGCTAT	GGCAACCCAC	AAGGACGACG	AAGAGCGATT
223_10	GTTAATCCCG	GTGTCGCCAT	GGCAACCCAC	AAGGACGACG	AGGAACGCTT
223_2	GTTAATCCCG	GTGTCGCCAT	GGCAACCCAC	AAGGACGACG	AGGAACGCTT
223_4	GTTAATCCCG	GTGTCGCCAT	GGCAACCCAC	AAGGACGACG	AGGAACGCTT
223_5	GTTAATCCCG	GTGTCGCCAT	GGCAACCCAC	AAGGACGACG	AGGAACGCTT
223_6	GTTAATCCCG	GTGTCGCCAT	GGCAACCCAC	AAGGACGACG	AGGAACGCTT
223_7	GTTAATCCCG	GTGTCGCCAT	GGCAACCCAC	AAGGACGACG	AGGAACGCTT
A3_4	GTCAATCCCG	GGCCCCCAAT	GGCCAGTCAC	AAGGACGATG	AGGAAAAGTA
A3_5	GTCAATCCCG	GGCCCCCAAT	GGCCAGTCAC	AAGGACGATG	AGGAAAAGTA
A3_7	GTCAATCCCG	GGCCCCCAAT	GGCCAGTCAC	AAGGACGATG	AGGAAAAGTA
A3_3	GTCAATCCCG	GGCCCCCAGT	GGCCAGTCAC	AAGGACGATG	AGGAAAAGTA
42_12	ACCAACCCCG	GCGTAGCCAT	GGCCACCAAC	AAGGACGACG	AGGACCAGTT
AAV1	ATCAACCCCTG	GCACTGCTAT	GGCCTCACAC	AAAGACGACG	AAGACAAGTT
AAV2	GTGAATCC..	GGCC...AT	GGCAAGCCAC	AAGGACGATG	AAGAAAAGTT
AAV3	GTGAATCCAG	GACCAGCTAT	GGCCAGTCAC	AAGGACGATG	AAGAAAAATT
AAV8	GCTAATCCTG	GCATCGCTAT	GGCAACACAC	AAAGACGACG	AGGAGCGTTT
AAV9	ATGAATCCTG	GCGTGGCTAT	GGCATCGCAC	AAAGACGACG	AGGACCGCTT
AAV7	GTTAATCCCG	GCGTCGCCAT	GGCAACTCAC	AAGGACGACG	AGGACCGCTT
44_2	GTAAATCCCG	GTGTCGCTAT	GGCAACCCAC	AAGGACGACG	AAGAGCGATT

Fig. 1AAZ

	3851					3900
42_2	CTTTCCCATC	AACGGAGTGC	TGGTTTTTTGG	CGAAACGGGG	GCTGCCAACA	
42_8	TTTTCCATCC	AGCGGAGTCT	TGATGTTTTGG	GAAACAGGGA	GCTGGAAA..	
42_15	TTTTCCATCC	AGCGGAGTCT	TGATGTTTTGG	GAAACAGGGA	GCTGGAAA..	
42_5b	TTTTCCATCC	AGCGGAGTCT	TGATGTTTTGG	GAAACAGGGA	GCTGGAAA..	
42_1b	TTTTCCATCC	AGCGGAGTCT	TGATGTTTTGG	GAAACAGGGA	GCTGGAAA..	
42_13	TTTTCCATCC	AGCGGAGTCT	TGATGTTTTGG	GAAACAGGGA	GCTGGAAA..	
42_3a	TTTTCCATCC	AGCGGAGTCT	TGATGTTTTGG	GAAACAGGGA	GCTGGAAA..	
42_4	TTTTCCATCC	AGCGGAGTCT	TGATGTTTTGG	GAAACAGGGA	GCTGGAAA..	
42_5a	TTTTCCATCC	AGCGGAGTCT	TGATGTTTTGG	GAAACAGGGA	GCTGGAAA..	
42_10	CTTTCCCATC	AACGGAGTGC	TGGTTTTTTGG	CAAAACGGGG	GCTGCCAACA	
42_3b	CTTTCCCATC	AACGGAGTGC	TGGTTTTTTGG	CAAAACGGGG	GCTGCCAACA	
42_11	CTTTCCCATC	AACGGAGTGC	TGGTTTTTTGG	CAAAACGGGG	GCTGCCAACA	
42_6b	CTTTCCCATC	AACGGAGTGC	TGGTTTTTTGG	CAAAACGGGG	GCTGCCAACA	
43_1	CTTCCCGTCA	AGCGGAGTTC	TAATGTTTTGG	CAAGCAGGGG	GCTGGAAA..	
43_5	CTTCCCGTCA	AGCGGAGTTC	TAATGTTTTGG	CAAGCAGGGG	GCTGGAAA..	
43_12	CTTCCCGTCA	AGCGGAGTTC	TAATGTTTTGG	CAAGCAGGGG	GCTGGAAA..	
43_20	CTTCCCTTCG	AGCGGGGTCC	TGATTTTTTGG	CAAGCAAGGA	GCCGGGAA..	
43_21	CTTCCCTTCG	AGCGGGGTCC	TGATTTTTTGG	CAAGCAAGGA	GCCGGGAA..	
43_23	CTTCCCTTCG	AGCGGGGTCC	TGATTTTTTGG	CAAGCAAGGA	GCCGGGAA..	
43_25	CTTCCCTTCG	AGCGGGGTCC	TGATTTTTTGG	CAAGCAAGGA	GCCGGGAA..	
44_1	TTTTCCGTCC	AGCGGAGTCT	TAATGTTTTGG	GAAACAGGGA	GCTGGAAA..	
44_5	TTTTCCGTCC	AGCGGAGTCT	TAATGTTTTGG	GAAACAGGGA	GCTGGAAA..	
223_10	CTTCCCTTCG	AGCGGAGTTC	TAATTTTTTGG	CAAAACTGGA	GCAGCTAATA	
223_2	CTCCCTTCG	AGCGGAGTTC	TAATTTTTTGG	CAAAACTGGA	GCAGCTAATA	
223_4	CTTCCCTTCG	AGCGGAGTTC	TAATTTTTTGG	CAAAACTGGA	GCAGCTAATA	
223_5	CTTCCCTTCG	AGCGGAGTTC	TAATTTTTTGG	CAAAACTGGA	GCAGCTAATA	
223_6	CTTCCCTTCG	AGCGGAGTTC	TAATTTTTTGG	CAAAACTGGA	GCAGCTAATA	
223_7	CTTCCCTTCG	AGCGGAGTTC	TAATTTTTTGG	CAAAACTGGA	GCAGCTAATA	
A3_4	TTTCCCCATG	CACGGAAATC	TCATCTTTTGG	AAAACAAGGC	ACAGGAAC..	
A3_5	TTTCCCCATG	CACGGAAATC	TCATCTTTTGG	AAAACAAGGC	ACAGGAAC..	
A3_7	TTTCCCCATG	CACGGAAATC	TCATCTTTTGG	AAAACAAGGC	ACAGGAAC..	
A3_3	TTTCCCCATG	CACGGAAATC	TCATCTTTTGG	AAAACAAGGC	ACAGGAAC..	
42_12	CTTTCCCATC	AACGGAGTGC	TGGTTTTTTGG	CAAAACGGGG	GCTGCCAACA	
AAV1	CTTTCCCATG	AGCGGTGTCA	TGATTTTTTGG	AAAAGAGAGC	GCCGGAGC..	
AAV2	TTTTCCCTCAG	AGCGGGGTTC	TCATCTTTTGG	GAAGCAAGGC	TCAGAGAA..	
AAV3	TTTCCCTATG	CACGGCAATC	TAATATTTTGG	CAAAGAAGGG	ACAACGGC..	
AAV8	TTTTCCCACT	AACGGGATCC	TGATTTTTTGG	CAAACAAAAT	GCTGCCAG..	
AAV9	CTTTCCCATCA	AGTGGCGTTC	TCATATTTTGG	CAAGCAAGGA	GCCGGGAA..	
AAV7	TTTCCCATCC	AGCGGAGTCC	TGATTTTTTGG	AAAAACTGGA	GCAACTAACA	
44_2	TTTTCCGTCC	AGCGGAGTCT	TAATGTTTTGG	GAAACAGGGA	GCTGGAAA..	

Fig. 1AAAA

	3901				3950
42_2	AGACAACGCT	GGAA.....	AACGTGCTAA	TGACCAGCGA	GGAGGAGATC
42_8	AGACAACG.T	GGACTATAGC	AGCGTTATGC	TAACCAGTGA	GGAAGAAATC
42_15	AGACAACG.T	GGACTATAGC	AGCGTTATGC	TAACCAGTGA	GGAAGAAATC
42_5b	AGACAACG.T	GGACTATAGC	AGCGTTATGC	TAACCAGTGA	GGAAGAAATC
42_1b	AGACAACG.T	GGACTATAGC	AGCGTTATGC	TAACCAGTGA	GGAAGAAATC
42_13	AGACAACG.T	GGACTATAGC	AGCGTTATGC	TAACCAGTGA	GGAAGAAATC
42_3a	AGACAACG.T	GGACTATAGC	AGCGTTATGC	TAACCAGTGA	GGAAGAAATC
42_4	AGACAACG.T	GGACTATAGC	AGCGTTATGC	TAACCAGTGA	GGAAGAAATC
42_5a	AGACAACG.T	GGACTATAGC	AGCGTTATGC	TAACCAGTGA	GGAAGAAATC
42_10	AGACAACGCT	GGAA.....	AACGTGCTAA	TGACCAGCGA	GGAGGAGATC
42_3b	AGACAACGCT	GGAA.....	AACGTGCTAA	TGACCAGCGA	GGAGGAGATC
42_11	AGACAACGCT	GGAA.....	AACGTGCTAA	TGACCAGCGA	GGAGGAGATC
42_6b	AGACAACGCT	GGAA.....	AACGTGCTAA	TGACCAGCGA	GGAGGAGATC
43_1	AGACAATG.T	GGACTACAGC	AGCGTGATGC	TCACCAGCGA	AGAAGAAATT
43_5	AGACAATG.T	GGACTACAGC	AGCGTGATGC	TCACCAGCGA	AGAAGAAATT
43_12	AGACAATG.T	GGACTACAGC	AGCGTGATGC	TCACCAGCGA	AGAAGAAATT
43_20	CGATGGAG.T	GGATTACAGC	CAAGTGCTGA	TTACAGATGA	GGAAGAAATC
43_21	CGATGGAG.T	GGATTACAGC	CAAGTGCTGA	TTACAGATGA	GGAAGAAATC
43_23	CGATGGAG.T	GGATTACAGC	CAAGTGCTGA	TTACAGATGA	GGAAGAAATC
43_25	CGATGGAG.T	GGATTACAGC	CAAGTGCTGA	TTACAGATGA	GGAAGAAATC
44_1	AGACAACG.T	GGACTATAGC	AGCGTTATGC	TAACCAGTGA	GGAAGAAATT
44_5	AGACAACG.T	GGACTATAGC	AGCGTTATGC	TAACCAGTGA	GGAAGAAATT
223_10	AAACTACATT	AGAA.....	AACGTGCTCA	TGACAAATGA	AGAAGAAATT
223_2	AAACTACATT	AGAA.....	AACGTGCTCA	TGACAAATGA	AGAAGAAATT
223_4	AAACTACATT	AGAA.....	AACGTGCTCA	TGACAAATGA	AGAAGAAATT
223_5	AAACTACATT	AGAA.....	AACGTGCTCA	TGACAAATGA	AGAAGAAATT
223_6	AAACTACATT	AGAA.....	AACGTGCTCA	TGACAAATGA	AGAAGAAATT
223_7	AAACTACATT	AGAA.....	AACGTGCTCA	TGACAAATGA	AGAAGAAATT
A3_4	TACCAATG.T	GGACATTGAA	TCAGTGCTTA	TTACAGACGA	AGAAGAAATC
A3_5	TACCAATG.T	GGACATTGAA	TCAGTGCTTA	TTACAGACGA	AGAAGAAATC
A3_7	TACCAATG.T	GGACATTGAA	TCAGTGCTTA	TTACAGACGA	AGAAGAAATC
A3_3	TACCAATG.T	GGACATTGAA	TCAGTGCTTA	TTACAGACGA	AGAAGAAATC
42_12	AGACAACGCT	GGAA.....	AACGTGCTAA	TGACCAGCGA	GGAGGAGATC
AAV1	TTCAAACA.C	TGCATTGGAC	AATGTCATGA	TTACAGACGA	AGAGGAAATT
AAV2	AACAAATG.T	GAACATTGAA	AAGGTCATGA	TTACAGACGA	AGAGGAAATC
AAV3	AAGTAACG.C	AGAATTAGAT	AATGTAATGA	TTACGGATGA	AGAAGAGATT
AAV8	AGACAATG.C	GGATTACAGC	GATGTCATGC	TCACCAGCGA	GGAAGAAATC
AAV9	CGATGGAG.T	CGACTACAGC	CAGGTGCTGA	TTACAGATGA	GGAAGAAATT
AAV7	AAACTACATT	GGAA.....	AATGTGTTAA	TGACAAATGA	AGAAGAAATT
44_2	AGACAACG.T	GGACTATAGC	AGCGTTATGC	TAACCAGTGA	GGAAGAAATT

Fig. 1AAAB

	3951				4000
42_2	AAAACCACCA	ATCCCGTGGC	TACAGAAGAA	TACGGTGTGG	TCTCCAGCAA
42_8	AAAACCACCA	ACCCAGTGGC	CACAGAACAG	TACGGCGTGG	TGGCCGATAA
42_15	AAAACCACCA	ACCCAGTGGC	CACAGAACAG	TACGGCGTGG	TGGCCGATAA
42_5b	AAAACCACCA	ACCCAGTGGC	CACAGAACAG	TACGGCGTGG	TGGCCGATAA
42_1b	AAAACCACCA	ACCCAGTGGC	CACAGAACAG	TACGGCGTGG	TGGCCGATAA
42_13	AAAACCACCA	ACCCAGTGGC	CACAGAACAG	TACGGCGTGG	TGGCCGATAA
42_3a	AAAACCACCA	ACCCAGTGGC	CACAGAACAG	TACGGCGTGG	TGGCCGATAA
42_4	AAAACCACCA	ACCCAGTGGC	CACAGAACAG	TACGGCGTGG	TGGCCGATAA
42_5a	AAAACCACCA	ACCCAGTGGC	CACAGAACAG	TACGGCGTGG	TGGCCGATAA
42_10	AAAACCACCA	ATCCCGTGGC	TACAGAAGAA	TACGGTGTGG	TCTCCAGCAA
42_3b	AAAACCACCA	ATCCCGTGGC	TACAGAACAG	TACGGTGTGG	TCTCCAGCAA
42_11	AAAACCACCA	ATCCCGTGGC	TACAGAAGAA	TACGGTGTGG	TCTCCAGCAA
42_6b	AAAACCACCA	ATCCCGTGGC	TACAGAAGAA	TACGGTGTGG	TCTCCAGCAA
43_1	AAAACCTACTA	ACCCAGTGGC	TACAGAGCAG	TATGGTGTGG	TGGCAGACAA
43_5	AAAACCTACTA	ACCCAGTGGC	TACAGAGCAG	TATGGTGTGG	TGGCAGACAA
43_12	AAAACCTACTA	ACCCAGTGGC	TACAGAGCAG	TATGGTGTGG	TGGCAGACAA
43_20	AAGGCTACCA	ACCCCGTGGC	CACAGAAGAA	TATGGAGCAG	TGGCCATCAA
43_21	AAGGCTACCA	ACCCCGTGGC	CACAGAAGAA	TATGGAGCAG	TGGCCATCAA
43_23	AAGGCTACCA	ACCCCGTGGC	CACAGAAGAA	TATGGAGCAG	TGGCCATCAA
43_25	AAGGCTACCA	ACCCCGTGGC	CACAGAAGAA	TATGGAGCAG	TGGCCATCAA
44_1	AAAACCACCA	ACCCAGTGGC	CACGGAACAG	TACGGCGTGG	TGGCCGATAA
44_5	AAAACCACCA	ACCCAGTGGC	CACAGAACAG	TACGGCGTGG	TGGCCGATAA
223_10	CGTCCTACCA	ACCCGGTAGC	TACCGAGGAA	TACGGGATTG	TAAGCAGCAA
223_2	CGTCCTACCA	ACCCGGTAGC	TACCGAGGAA	TACGGGATTG	TAAGCAGCAA
223_4	CGTCCTACCA	ACCCGGTAGC	TACCGAGGAA	TACGGGATTG	TAAGCAGCAA
223_5	CGTCCTACCA	ACCCGGTAGC	TACCGAGGAA	TACGGGATTG	TAAGCAGCAA
223_6	CGTCCTACCA	ACCCGGTAGC	TACCGAGGAA	TACGGGATTG	TAAGCAGCAA
223_7	CGTCCTACCA	ACCCGGTAGC	TACCGAGGAA	TACGGGATTG	TAAGCAGCAA
A3_4	AGAACAACCTA	ATCCTGTGGC	TACAGAACAA	TACGGACAGG	TTGCCACCAA
A3_5	AGAACGACTA	ATCCTGTGGC	TACAGAACAA	TACGGACAGG	TTGCCACCAA
A3_7	AGAACAACCTA	ATCCTGTGGC	TACAGAACAA	TACGGACAGG	TTGCCACCAA
A3_3	AGAACAACCTA	ATCCTGTGGC	TACAGAACAA	TACGGACAGG	TTGCCACCAA
42_12	AAAACCACCA	ATCCCGTGGC	TACAGAAGAA	TACGGTGTGG	TCTCCAGCAA
AAV1	AAAGCCACTA	ACCCTGTGGC	CACCGAAAGA	TTTGGGACCG	TGGCAGTCAA
AAV2	GGAACAACCA	ATCCCGTGGC	TACGGAGCAG	TATGGTTCTG	TATCTACCAA
AAV3	CGTACCACCA	ATCCTGTGGC	AACAGAGCAG	TATGGAACTG	TGGCAAATAA
AAV8	AAAACCACCTA	ACCCTGTGGC	TACAGAGGAA	TACGGTATCG	TGGCAGATAA
AAV9	AAAGCCACCA	ACCCTGTAGC	CACAGAGGAA	TACGGAGCAG	TGGCCATCAA
AAV7	CGTCCTACTA	ATCCTGTAGC	CACGGAAGAA	TACGGGATAG	TCAGCAGCAA
44_2	AAAACCACCA	ACCCAGTGGC	CACAGAACAG	TACGGCGTGG	TGGCCGATAA

Fig. 1AAAC

	4001				4050
42_2	CCTGCAATCG	TCTACGGCCG	GACCCCAGAC	ACAGACTGTC	AACAGCCAGG
42_8	CCTGCAACAG	CAAAACGCCG	CTCCTATTGT	AGGGGCCGTC	AACAGTCAAG
42_15	CCTGCAACAG	CAAAACGCCG	CTCCTATTGT	AGGGGCCGTC	AACAGTCAAG
42_5b	CCTGCAACAG	CAAAACGCCG	CTCCTATTGT	AGGGGCCGTC	AACAGTCAAG
42_1b	CCTGCAACAG	CAAAACGCCG	CTCCTATTGT	AGGGGCCGTC	AACAGTCAAG
42_13	CCTGCAACAG	CAAAACGCCG	CTCCTATTGT	AGGGGCCGTC	AACAGTCAAG
42_3a	CCTGCAACAG	CAAAACGCCG	CTCCTATTGT	AGGGGCCGTC	AACAGTCAAG
42_4	CCTGCAACAG	CAAAACGCCG	CTCCTATTGT	AGGGGCCGTC	AACAGTCAAG
42_5a	CCTGCAACAG	CAAAACGCCG	CTCCTATTGT	AGGGGCCGTC	AACAGTCAAG
42_10	CCTGCAATCG	TCTACGGCCG	GACCCCAGAC	ACAGACTGTC	AACAGCCAGG
42_3b	CCTGCAATCG	TCTACGGCCG	GACCCCAGAC	ACAGACTGTC	AACAGCCAGG
42_11	CCTGCAATCG	TCTACGGCCG	GACCCCAGAC	ACAGACTGTC	AACAGCCAGG
42_6b	CCTGCAATCG	TCTACGGCCG	GACCCCAGAC	ACAGACTGTC	AACAGCCAGG
43_1	CCTGCAGCAG	ACCAACGGAG	CTCCCATTTGT	GGGAACTGTC	AACAGCCAGG
43_5	CCTGCAGCAG	ACCAACGGAG	CTCCCATTTGT	GGGAACTGTC	AACAGCCAGG
43_12	CCTGCAGCAG	ACCAACGGAG	CTCCCATTTGT	GGGAACTGTC	AACAGCCAGG
43_20	CAACCAGGCC	GCCAATACGC	AGGCGCAGAC	CGGACTCGTG	CACAACCAGG
43_21	CAACCAGGCC	GCCAATACGC	AGGCGCAGAC	CGGACTCGTG	CACAACCAGG
43_23	CAACCAGGCC	GCCAATACGC	AGGCGCAGAC	CGGACTCGTG	CACAACCAGG
43_25	CAACCAGGCC	GCCAATACGC	AGGCGCAGAC	CGGACTCGTG	CACAACCAGG
44_1	CCTGCAACAG	CAAAACGCCG	CTCCTATTGT	AGGGGCCGTC	AACAGTCAAG
44_5	CCTGCAACAG	CAAAACGCCG	CTCCTATTGT	AGGGGCCGTC	AACAGTCAAG
223_10	CTTGCAAGCG	GCTAGCACCG	CAGCCCAGAC	ACAAGTTGTT	AACAACCAGG
223_2	CTTGCAAGCG	GCTAGCACCG	CAGCCCAGAC	ACAAGTTGTT	AACAACCAGG
223_4	CTTGCAAGCG	GCTAGCACCG	CAGCCCAGAC	ACAAGTTGTT	AACAACCAGG
223_5	CTTGCAAGCG	GCTAGCACCG	CAGCCCAGAC	ACAAGTTGTT	AACAACCAGG
223_6	CTTGCAAGCG	GCTAGCACCG	CAGCCCAGAC	ACAAGTTGTT	AACAACCAGG
223_7	CTTGCAAGCG	GCTAGCACCG	CAGCCCAGAC	ACAAGTTGTT	AACAACCAGG
A3_4	CCATCAGAGT	CAGGACACCA	CAGCTTCCTA	TGGAAGTGTG	GACAGCCAGG
A3_5	CCGTCAGAGT	CAGAACACCA	CAGCTTCCTA	TGGAAGTGTG	GACAGCCAGG
A3_7	CCATCAGAGT	CAGAACACCA	CAGCTTCCTA	TGGAAGTGTG	GACAGCCAGG
A3_3	CCATCAGAGT	CAGAACACCA	CAGCTTCCTA	TGGAAGTGTG	GACAGCCAGG
42_12	CCTGCAATCG	TCTACGGCCG	GACCCCAGAC	ACAGACTGTC	AACAGCCAGG
AAV1	TTTCCAGAGC	AGCAGCACAG	ACCCTGCGAC	CGGAGATGTC	CATGCTATGG
AAV2	CCTCCAGAGA	GGCAACAGAC	AAGCAGCTAC	CGCAGATGTC	AACACACAAG
AAV3	CTTGCAAGCG	TCAAATACAG	CTCCACGAC	TGGAAGTGTG	AATCATCAGG
AAV8	CTTGCAAGCG	CAAAACACGG	CTCCTCAAAT	TGGAAGTGTG	AACAGCCAGG
AAV9	CAACCAGGCC	GCTAACACGC	AGGCGCAAAC	TGGACTTGTG	CATAACCAGG
AAV7	CTTACAAGCG	GCTAATACTG	CAGCCCAGAC	ACAAGTTGTC	AACAACCAGG
44_2	CCTGCAACAG	CAAAACGCCG	CTCCTATTGT	AGGGGCCGTC	AACAGTCAAG

Fig. 1AAAD

	4051				4100
42_2	GGGCTCTGCC	CGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
42_8	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
42_15	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
42_5b	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
42_1b	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
42_13	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
42_3a	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
42_4	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
42_5a	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
42_10	GGGCTCTGCC	CGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
42_3b	GGGCTCTGCC	CGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
42_11	GGGCTCTGCC	CGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
42_6b	GGGCTCTGCC	CGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
43_1	GGGCCTTACC	TGGTATGGTC	TGGCAAAACC	GGGACGTGTA	CCTGCAGGGC
43_5	GGGCCTTACC	TGGTATGGTC	TGGCAAAACC	GGGACGTGTA	CCTGCAGGGC
43_12	GGGCCTTACC	TGGTATGGTC	TGGCAAAACC	GGGACGTGTA	CCTGCAGGGC
43_20	GGGTGATTCC	CGGCATGGTG	TGGCAGAATA	GAGACGTGTA	CCTGCAGGGT
43_21	GGGTGATTCC	CGGCATGGTG	TGGCAGAATA	GAGACGTGTA	CCTGCAGGGT
43_23	GGGTGATTCC	CGGCATGGTG	TGGCAGAATA	GAGACGTGTA	CCTGCAGGGT
43_25	GGGTGATTCC	CGGCATGGTG	TGGCAGAATA	GAGACGTGTA	CCTGCAGGGT
44_1	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
44_5	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
223_10	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAAGGT
223_2	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAAGGT
223_4	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAAGGT
223_5	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAAGGT
223_6	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAAGGT
223_7	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAAGGT
A3_4	GAATCTTACC	TGGAATGGTG	TGGCAGGACC	GCGATGTCTA	TCTTCAAGGT
A3_5	GAATCTTACC	TGGAATGGTG	TGGCAGGACC	GCGATGTCTA	TCTTCAAGGT
A3_7	GAATCTTACC	TGGAATGGTG	TGGCAGGACC	GCGATGTCTA	TCTTCAAGGT
A3_3	GAATCTTACC	TGGAATGGTG	TGGCAGGACC	GCGATGTCTA	TCTTCAAGGT
42_12	GGGCTCTGCC	CGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
AAV1	GAGCATTACC	TGGCATGGTG	TGGCAAGATA	GAGACGTGTA	CCTGCAGGGT
AAV2	GCGTTCTTCC	AGGCATGGTC	TGGCAGGACA	GAGATGTGTA	CCTTCAGGGG
AAV3	GGGCCTTACC	TGGCATGGTG	TGGCAAGATC	GTGACGTGTA	CCTTCAGGGA
AAV8	GGGCCTTACC	CGGTATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
AAV9	GAGTTATTCC	TGGTATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGC
AAV7	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT
44_2	GAGCCTTACC	TGGCATGGTC	TGGCAGAACC	GGGACGTGTA	CCTGCAGGGT

Fig. 1AAAE

	4101				4150
42_2	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
42_8	CCT.ATCTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
42_15	CCT.ATCTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
42_5b	CCT.ATCTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
42_1b	CCT.ATCTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
42_13	CCT.ATCTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
42_3a	CCT.ATCTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
42_4	CCT.ATCTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
42_5a	CCT.ATCTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
42_10	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
42_3b	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
42_11	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
42_6b	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
43_1	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
43_5	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
43_12	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
43_20	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
43_21	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
43_23	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
43_25	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
44_1	CCT.ATCTGG	GCCAAGATTC	CTCACACGGA	CGGAAACTTT	CATCCCTCGC
44_5	CCT.ATCTGG	GCCAAGATTC	CTCACACGGA	CGGAAACTTT	CATCCCTCGC
223_10	CCC.ATTTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
223_2	CCC.ATTTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
223_4	CCC.ATTTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
223_5	CCC.ATTTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
223_6	CCC.ATTTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
223_7	CCC.ATTTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
A3_4	CCC.ATTTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
A3_5	CCC.ATTTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
A3_7	CCC.ATTTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
A3_3	CCC.ATTTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CATCCTTCGC
42_12	CCC.ATCTGG	GCCAAAATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
AAV1	CCC.ATTTGG	GCCAAAATTC	CTCACACGGA	TGGCAACTTT	CACCCGTCTC
AAV2	CCC.ATCTGG	GCAAAGATTC	CACACACGGA	CGGCAACTTT	CACCCGTCTC
AAV3	CCT.ATCTGG	GCAAAGATTC	CTCACACGGA	TGGCAACTTT	CATCCTTCGC
AAV8	CCC.ATCTGG	GCCAAGATTC	CTCACACGGA	CGGCAACTTT	CACCCGTCTC
AAV9	CCCTATTTGG	GCTAAAATAC	CTCACACGGA	TGGCAACTTT	CACCCGTCTC
AAV7	CCC.ATCTGG	GCCAAGATTC	CTCACACGGA	TGGCAACTTT	CACCCGTCTC
44_2	CCT.ATCTGG	GCCAAGATTC	CTCACACGGA	CGGAAACTTT	CATCCCTCGC

Fig. 1AAAF

	4151				4200
42_2	CCCTGATGGG	CGGATTTGGA	CTCAAACACC	CGCCTCCTCA	AATTCTCATC
42_8	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGATT
42_15	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGATT
42_5b	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGATT
42_1b	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGATT
42_13	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGATT
42_3a	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGATT
42_4	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGATT
42_5a	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGATT
42_10	CCCTGATGGG	CGGATTTGGA	CTCAAACACC	CGCCTCCTCA	AATTCTCATC
42_3b	CCCTGATGGG	CGGATTTGGA	CTCAAACACC	CGCCTCCTCA	AATTCTCATC
42_11	CCCTGATGGG	CGGATTTGGA	CTCAAACACC	CGCCTCCTCA	AATTCTCATC
42_6b	CCCTGATGGA	CGGATTTGGA	CTCAAACACC	CGCCTCCTCA	AATTCTCATC
43_1	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGGTG
43_5	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGGTG
43_12	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGGTG
43_20	CCCTGATGGG	CGGCTTTGGA	CTGAAGCACC	CGCCTCCTCA	AATTCTCATC
43_21	CCCTGATGGG	CGGCTTTGGA	CTGAAGCACC	CGCCTCCTCA	AATTCTCATC
43_23	CCCTGATGGG	CGGCTTTGGA	CTGAAGCACC	CGCCTCCTCA	AATTCTCATC
43_25	CCCTGATGGG	CGGCTTTGGA	CTGAAGCACC	CGCCTCCTCA	AATTCTCATC
44_1	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGATT
44_5	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGATT
223_10	CTCTAATGGG	TGGCTTTGGA	CTGAAACACC	CGCCTCCCCA	GATCCTGATC
223_2	CTCTAATGGG	TGGCTTTGGA	CTGAAACACC	CGCCTCCCCA	GATCCTGATC
223_4	CTCTAATGGG	TGGCTTTGGA	CTGAAACACC	CGCCTCCCCA	GATCCTGATC
223_5	CTCTAATGGG	TGGCTTTGGA	CTGAAACACC	CGCCTCCCCA	GATCCTGATC
223_6	CTCTAATGGG	TGGCTTTGGA	CTGAAACACC	CGCCTCCCCA	GATCCTGATC
223_7	CTCTAATGGG	TGGCTTTGGA	CTGAAACACC	CGCCTCCCCA	GATCCTGATC
A3_4	CGCTCATGGG	AGGCTTTGGA	CTGAAACACC	CTCCTCCCCA	GATCCTGATC
A3_5	CGCTCATGGG	AGGCTTTGGA	CTGAAACACC	CTCCTCCCCA	GATCCTGATC
A3_7	CGCTCATGGG	AGGCTTTGGA	CTGAAACACC	CTCCTCCCCA	GATCCTGATC
A3_3	CGCTCATGGG	AGGCTTTGGA	CTGAAACACC	CTCCTCCCCA	GATCCTGATC
42_12	CCCTGATGGG	CGGATTTGGA	CTCAAACACC	CGCCTCCTCA	AATTCTCATC
AAV1	CTCTTATGGG	CGGCTTTGGA	CTCAAGAACC	CGCCTCCTCA	GATCCTCATC
AAV2	CCCTCATGGG	TGGATTCGGA	CTTAAACACC	CTCCTCCACA	GATTCTCATC
AAV3	CTCTGATGGG	AGGCTTTGGA	CTGAAACATC	CGCCTCCTCA	AATCATGATC
AAV8	CGCTGATGGG	CGGCTTTGGC	CTGAAACATC	CTCCGCCTCA	GATCCTGATC
AAV9	CTCTGATGGG	TGGATTTGGA	CTGAAACACC	CACCTCCACA	GATTCTAATT
AAV7	CTTTGATGGG	CGGCTTTGGA	CTTAAACATC	CGCCTCCTCA	GATCCTGATC
44_2	CGCTGATGGG	AGGCTTTGGA	CTGAAACACC	CGCCTCCTCA	GATCCTGATT

Fig. 1AAAG

	4201				4250
42_2	AAAAACACCC	CGGTACCTGC	TAATCCTCCA	GAGGTGTTTA	CTCCTGCCAA
42_8	AAGAATACAC	CTGTTCCCGC	GGATCCTCCA	ACTACCTTCA	GTCAAGCCAA
42_15	AAGAATACAC	CTGTTCCCGC	GGATCCTCCA	ACTACCTTCA	GTCAAGCCAA
42_5b	AAGAATACAC	CTGTTCCCGC	GGATCCTCCA	ACTACCTTCA	GTCAAGCCAA
42_1b	AAGAATACAC	CTGTTCCCGC	GGATCCTCCA	ACTACCTTCA	GTCAAGCCAA
42_13	AAGAATACAC	CTGTTCCCGC	GGATCCTCCA	ACTACCTTCA	GTCAAGCCAA
42_3a	AAGAATACAC	CTGTTCCCGC	GGATCCTCCA	ACTACCTTCA	GTCAAGCCAA
42_4	AAGAATACAC	CTGTTCCCGC	GGATCCTCCA	ACTACCTTCA	GTCAAGCCAA
42_5a	AAGAATACAC	CTGTTCCCGC	GGATCCTCCA	ACTACCTTCA	GTCAAGCCAA
42_10	AAAAACACCC	CGGTACCTGC	TAATCCTCCA	GAGGTGTTTA	CTCCTGCCAA
42_3b	AAAAACACCC	CGGTACCTGC	TAATCCTCCA	GAGGTGTTTA	CTCCTGCCAA
42_11	AAAAACACCC	CGGTACCTGC	TAATCCTCCA	GAGGTGTTTA	CTCCTGCCAA
42_6b	AAAAACACCC	CGGTACCTGC	TAATCCTCCA	GAGGTGTTTA	CTCCTGCCAA
43_1	AAAAACACTC	CTGTTCCCTGC	GGATCCTCCG	ACCACCTTCA	GCCAGGCCAA
43_5	AAAAACACTC	CTGTTCCCTGC	GGATCCTCCG	ACCACCTTCA	GCCAGGCCAA
43_12	AAAAACACTC	CTGTTCCCTGC	GGATCCTCCG	ACCACCTTCA	GCCAGGCCAA
43_20	AAGAACACAC	CGGTTCCAGC	GGACCCGCCG	CTTACCTTCA	ACCAGGCCAA
43_21	AAGAACACAC	CGGTTCCAGC	GGACCCGCCG	CTTACCTTCA	ACCAGGCCAA
43_23	AAGAACACAC	CGGTTCCAGC	GGACCCGCCG	CTTACCTTCA	ACCAGGCCAA
43_25	AAGAACACAC	CGGTTCCAGC	GGACCCGCCG	CTTACCTTCA	ACCAGGCCAA
44_1	AAGAATACAC	CTGTTCCCGC	GGATCCTCCA	ACTACCTTCA	GTCAAGCTAA
44_5	AAGAATACAC	CTGTTCCCGC	GGATCCTCCA	ACTACCTTCA	GTCAAGCTAA
223_10	AAAAACACAC	CGGTACCTGC	TAATCCTCCA	GAAGTGTTTA	CTCCTGCCAA
223_2	AAAAACACGC	CGGTACCTGC	TAATCCTCCA	GAAGTGTTTA	CTCCTGCCAA
223_4	AAAAACACAC	CGGTACCTGC	TAATCCTCCA	GAAGTGTTTA	CTCCTGCCAA
223_5	AAAAACACAC	CGGTACCTGC	TAATCCTCCA	GAAGTGTTTA	CTCCTGCCAA
223_6	AAAAACACAC	CGGTACCTGC	TAATCCTCCA	GAAGTGTTTA	CTCCTGCCAA
223_7	AAAAACACAC	CGGTACCTGC	TAATCCTCCA	GAAGTGTTTA	CTCCTGCCAA
A3_4	AAAAACACAC	CTGTGCCAGC	GAATCCCGCG	ACCACTTTCA	CTCCTGGAAA
A3_5	AAAAACACAC	CTGTGCCAGC	GAATCCCGCG	ACCACTTTCA	CTCCTGGAAA
A3_7	AAAAACACAC	CTGTGCCAGC	GAATCCCGCG	ACCACTTTCA	CTCCTGGAAA
A3_3	AAAAACACAC	CTGTGCCAGC	GAATCCCGCG	ACCACTTTCA	CTCCTGGAAA
42_12	A...A.....	.....	.....	.....	.....
AAV1	AAAAACACGC	CTGTTCCCTGC	GAATCCTCCG	GCGGAGTTTT	CAGCTACAAA
AAV2	AAGAACACCC	CGGTACCTGC	GAATCCTTCG	ACCACCTTCA	GTGCGGCCAA
AAV3	AAAAATACTC	CGGTACCGGC	AAATCCTCCG	ACGACTTTCA	GCCCGGCCAA
AAV8	AAGAACACGC	CTGTACCTGC	GGATCCTCCG	ACCACCTTCA	ACCAGTCAAA
AAV9	AAAAATACAC	CAGTGCCGGC	AGATCCTCCT	CTTACCTTCA	ATCAAGCCAA
AAV7	AAGAACACTC	CCGTTCCCGC	TAATCCTCCG	GAGGTGTTTA	CTCCTGCCAA
44_2	AAGAATACAC	CTGTTCCCGC	GGATCCTCCA	ACTACCTTCA	GTCAAGCTAA

Fig. 1AAAH

	4251				4300
42_2	GTTTGCCTCA	TTTATCACGC	AGTACAGCAC	CGGCCA.GGT	CAGCGTGGAG
42_8	GCTGGCGTCG	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
42_15	GCTGGCGTCG	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
42_5b	GCTGGCGTCG	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
42_1b	GCTGGCGTCG	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
42_13	GCTGGCGTCG	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
42_3a	GCTGGCGTCG	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
42_4	GCCGGCGTCG	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
42_5a	GCTGGCGTCG	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
42_10	GTTTGCCTCA	TTTATCACGC	AGTACAGCAC	CGGCCA.GGT	CAGCGTGGAG
42_3b	GTTTGCCTCA	TTTATCACGC	AGTACAGCAC	CGGCCA.GGT	CAGCGTGGAG
42_11	GTTTGCCTCA	TTTATCACGC	AGTACAGCAC	CGGCCA.GGT	CAGCGTGGAG
42_6b	GTTTGCCTCA	TTTATCACGC	AGTACAGCAC	CGGCCA.GGT	CAGCGTGGAG
43_1	GCTGGCTTCT	TTTATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
43_5	GCTGGCTTCT	TTTATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
43_12	GCTGGCTTCT	TTTATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
43_20	GCTGAACTCT	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
43_21	GCTGAACTCT	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
43_23	GCTGAACTCT	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
43_25	GCTGAACTCT	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
44_1	GCTGGCGTCG	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
44_5	GCTGGCGTCG	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA
223_10	GTTTGCCTTCC	TTCATCACGC	AGTACAGCAC	CGGGCA.AGT	CAGCGTTGAG
223_2	GTTTGCCTTCC	TTCATCACGC	AGTACAGCAC	CGGGCA.AGT	CAGCGTTGAG
223_4	GTTTGCCTTCC	TTCATCACGC	AGTACAGCAC	CGGGCA.AGT	CAGCGTTGAG
223_5	GTTTGCCTTCC	TTCATCACGC	AGTACAGCAC	CGGGCA.AGT	CAGCGTTGAG
223_6	GCTTGCTTCC	TTCATCACGC	AGTACAGCAC	CGGGCA.AGT	CAGCGTTGAG
223_7	GATTGCTTCC	TTCATCACGC	AGTACAGCAC	CGGGCA.AGT	CAGCGTTGAG
A3_4	GTTTGCCTTCG	TTCATTACCC	AGTATTCCAC	CGGACA.GGT	CAGCGTGGAA
A3_5	GTTTGCCTTCG	TTCATTACCC	AGTATTCCAC	CGGACA.GGT	CAGCGTGGAA
A3_7	GTTTGCCTTCG	TTCATTACCC	AGTATTCCAC	CGGACA.GGT	CAGCGTGGAA
A3_3	GTTTGCCTTCG	TTCATTACCC	AGTATTCCAC	CGGACA.GGT	CAGCGTGGAA
42_12	.....	.....	.....	.....	.....
AAV1	GTTTGCCTTCA	TTCATCACCC	AATACTCCAC	AGGACA.AGT	GAGTGTGGAA
AAV2	GTTTGCCTTCC	TTCATCACAC	AGTACTCCAC	GGGACACGGT	CAGCGTGGAG
AAV3	GTTTGCCTTCA	TTTATCACTC	AGTACTCCAC	TGGACA.GGT	CAGCGTGGAA
AAV8	GCTGAACTCT	TTCATCACGC	AATACAGCAC	CGGACA.GGT	CAGCGTGGAA
AAV9	GCTGAACTCT	TTCATCACGC	AGTACAGCAC	GGGACA.AGT	CAGCGTGGAA
AAV7	GTTTGCCTTCG	TTCATCACAC	AGTACAGCAC	CGGACA.AGT	CAGCGTGGAA
44_2	GCTGGCGTCG	TTCATCACGC	AGTACAGCAC	CGGACA.GGT	CAGCGTGGAA

Fig. 1AAAI

	4301				4350
42_2	ATCGAGTGGG	AACTGCAGAA	AGAAAACAGC	AAACGCTGGA	ATCCAGAGAT
42_8	ATTGAATGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
42_15	ATTGAATGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
42_5b	ATTGAATGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
42_1b	ATTGAATGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
42_13	ATTGAATGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
42_3a	ATTGAATGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
42_4	ATTGAATGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
42_5a	ATTGAATGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
42_10	ATCGAGTGGG	AACTGCAGAA	AGAAAACAGC	AAACGCTGGA	ATCCAGAGAT
42_3b	ATCGAGTGGG	AACTGCAGAA	AGAAAACAGC	AAACGCTGGA	ATCCAGAGAT
42_11	ATCGAGTGGG	AACTGCAGAA	AGAGAACAGC	AAACGCTGGA	ATCCAGAGAT
42_6b	ATCGAGTGGG	AACTGCAGAA	AGAAAACAGC	AAACGCTGGA	ATCCAGAGAT
43_1	ATCGAATGGG	AGCTGCAGAA	AGAAAACAGC	AAGCGCTGGA	ACCCAGAGAT
43_5	ATCGAATGGG	AGCTGCAGAA	AGAAAACAGC	AAGCGCTGGA	ACCCAGAGAT
43_12	ATCGAATGGG	AGCTGCAGAA	AGAAAACAGC	AAGCGCTGGA	ACCCAGAGAT
43_20	ATCGAGTGGG	AGCTGCAGAA	AGAAAACAGC	AAACGCTGGA	ATCCAGAGAT
43_21	ATCGAGTGGG	AGCTGCAGAA	AGAAAACAGC	AAACGCTGGA	ATCCAGAGAT
43_23	ATCGAGTGGG	AGCTGCAGAA	AGAAAACAGC	AAACGCTGGA	ATCCAGAGAT
43_25	ATCGAGTGGG	AGCTGCAGAA	AGAAAACAGC	AAACGCTGGA	ATCCAGAGAT
44_1	ATTGAATGGG	AGCTGCAGAA	AGAAAACAGC	AAACGCTGGA	ACCCAGAGAT
44_5	ATTGAATGGG	AGCTGCAGAA	AGAAAACAGC	AAACGCTGGA	ACCCAGAGAT
223_10	ATCGAGTGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
223_2	ATCGAGTGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
223_4	ATCGAATGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
223_5	ATCGAATGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
223_6	ATCGAGTGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
223_7	ATCGAGTGGG	AGCTGCAGAA	AGAGAACAGC	AAGCGCTGGA	ACCCAGAGAT
A3_4	ATAGAGTGGG	AGCTGCAGAA	AGAAAACAGC	AAACGCTGGA	ACCCAGAAAT
A3_5	ATAGAGTGGG	AGCTGCAGAA	AGAAAACAGC	AAACGCTGGA	ACCCGGAAT
A3_7	ATAGAGTGGG	AGCTGCAGAA	AGAAAACAGC	AAACGCTGGA	ACCCAGAAAT
A3_3	ATAGAGTGGG	AGCTGCAGAA	AGAAAACAGC	AAACGCTGGA	ACCCAGAAAT
42_12	.....	.....	.....	.....	.....
AAV1	ATTGAATGGG	AGCTGCAGAA	AGAAAACAGC	AAGCGCTGGA	ATCCCGAAGT
AAV2	ATCGAGTGGG	AGCTGCAGAA	GGAAAACAGC	AAACGCTGGA	ATCCCGAAT
AAV3	ATTGAGTGGG	AGCTACAGAA	AGAAAACAGC	AAACGTTGGA	ATCCAGAGAT
AAV8	ATTGAATGGG	AGCTGCAGAA	GGAAAACAGC	AAGCGCTGGA	ACCCGAGAT
AAV9	ATCGAGTGGG	AGCTGCAGAA	AGAAAACAGC	AAGCGCTGGA	ATCCAGAGAT
AAV7	ATCGAGTGGG	AGCTGCAGAA	GGAAAACAGC	AAGCGCTGGA	ACCCGAGAT
44_2	ATTGAATGGG	AGCTGCAGAA	AGAAAACAGC	AAACGCTGGA	ACCCAGAGAT

Fig. 1AAAJ

	4351				4400
42_2	TCAGTACACC	TCAAATTATG	CCAAGTCTAA	TAAT.GTGGA	ATTTGCTGTC
42_8	TCAGTATACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
42_15	TCAGTATACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
42_5b	TCAGTATACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
42_1b	TCAGTATACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
42_13	TCAGTATACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
42_3a	TCAGTATACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
42_4	TCAGTATACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
42_5a	TCAGTATACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
42_10	TCAGTACACC	TCAAATTATG	CCAAGTCTAA	TAAT.GTGGA	ATTTGCTGTC
42_3b	TCAGTACACC	TCAAATTATG	CCAAGTCTAA	TAAT.GTGGA	ATTTGCTGTC
42_11	TCAGTACACC	TCAAATTATG	CCAAGTCTAA	TAAT.GTGGA	ATTTGCTGTC
42_6b	TCAGTACACC	TCAAATTATG	CCAAGTCTAA	TAAT.GTGGA	ATTTGCTGTC
43_1	TCAGTATACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
43_5	TCAGTATACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
43_12	TCAGTATACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
43_20	TCAATACACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
43_21	TCAATACACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
43_23	TCAATACACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
43_25	TCAATACACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
44_1	TCAATACACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTCGCTGTT
44_5	TCAATACACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTT
223_10	TCAGTACACC	TCCAACCTTTG	ACAAACAGAC	TGGA.GTGGA	CTTTGCTGTT
223_2	TCAGTACACC	TCCAACCTTTG	ACAAACAGAC	TGGA.GTGGA	CTTTGCTGTT
223_4	TCAGTACACC	TCCAACCTTTG	ACAAACAGAC	TGGA.GTGGA	CTTTGCTGTT
223_5	TCAGTACACC	TCCAACCTTTG	ACAAACAGAC	TGGA.GTGGA	CTTTGCTGTT
223_6	TCAGTACACC	TCCAACCTTTG	ACAAACAGAC	TGGA.GTGGA	CTTTGCTGTT
223_7	TCAGTACACC	TCCAACCTTTG	ACAAACAGAC	TGGA.GTGGA	CTTTGCTGTT
A3_4	TCAGTACACC	TCCAACACTACA	ACAAGTCGGT	GAAT.GTGGA	GTTTACCGTG
A3_5	TCAGTACACC	TCCAACACTACA	ACAAGTCGGT	GAAT.GTGGA	GTTTACCGTG
A3_7	TCAGTACACC	TCCAACACTACA	ACAAGTCGGT	GAAT.GTGGA	GTTTACCGTG
A3_3	TCAGTACACC	TCCAACACTACA	ACAAGTCGGT	GAAT.GTGGA	GTTTACCGTG
42_12	...GTATACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
AAV1	GCAGTACACA	TCCAATTATG	CAAAATCTGC	CAAC.GTTGA	TTTTACTGTG
AAV2	TCAGTACACT	TCCAACACTACA	ACAAGTCTGT	TAATCGTGGA	CTT.ACCGTG
AAV3	TCAGTACACT	TCCAACACTACA	ACAAGTCTGT	TAAT.GTGGA	CTTTACTGTA
AAV8	CCAGTACACC	TCCAACACTACT	ACAAATCTAC	AAGT.GTGGA	CTTTGCTGTT
AAV9	CCAGTATACT	TCAAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTC
AAV7	TCAGTACACC	TCCAACCTTTG	AAAAGCAGAC	TGGT.GTGGA	CTTTGCCGTT
44_2	TCAATACACT	TCCAACACTACT	ACAAATCTAC	AAAT.GTGGA	CTTTGCTGTT

Fig. 1AAAK

	4401					4450
42_2	AACAACGAAG	GGGTTTATAC	TGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
42_8	AATACTGAGG	GTACTTATTC	AGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
42_15	AATACTGAGG	GTACTTATTC	AGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
42_5b	AATACTGAGG	GTACTTATTC	AGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
42_1b	AATACTGAGG	GTACTTATTC	AGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
42_13	AATACTGAGG	GTACTTATTC	AGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
42_3a	AATACTGAGG	GTACTTATTC	AGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
42_4	AATACTGAGG	GTACTTATTC	AGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
42_5a	AATACTGAGG	GTACTTATTC	AGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
42_10	AACAACGAAG	GGGTTTATAC	TGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
42_3b	AACAACGAAG	GGGTTTATAC	TGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
42_11	AACAACGAAG	GGGTTTATAC	TGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
42_6b	AACAACGAAG	GGGTTTATAC	TGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
43_1	AATACTGAGG	GTACTTATTC	AGAGCCTCGC	CCCATTGGCA	CTCGTTATCT	
43_5	AATACCGAGG	GTACTTATTC	AGAGCCTCGC	CCCATTGGCA	CTCGTTATCT	
43_12	AATACTGAGG	GTACTTATTC	AGAGCCTCGC	CCCATTGGCA	CTCGTTATCT	
43_20	AACACGGAAG	GAGTTTATAG	CGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
43_21	AACACGGAAG	GAGTTTATAG	CGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
43_23	AACACGGAAG	GAGTTTATAG	CGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
43_25	AACACGGAGG	GGGTTTATAG	CGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
44_1	AACACAGATG	GCACTTATTC	TGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
44_5	AACACAGATG	GCACTTATTC	TGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
223_10	GACAGCCAGG	GTGTTTACTC	TGAGCCT...	.....	.....	
223_2	GACAGCCAGG	GTGTTTACTC	TGAGCCT...	.....	.....	
223_4	GACAGCCAGG	GTGTTTACTC	TGAGCCT...	.....	.....	
223_5	GACAGCCAGG	GTGTTTACTC	TGAGCCT...	.....	.....	
223_6	GACAGCCAGG	GTGTTTACTC	TGAGCCT...	.....	.....	
223_7	GACAGCCAGG	GTGTTTACTC	TGAGCCT...	.....	.....	
A3_4	GACGCAAACG	GTGTTTATTC	TGAACCCCGC	CCTATTGGCA	CTCGTTACCT	
A3_5	GACGCAAACG	GTGTTTATTC	TGAACCCCGC	CCTATTGGCA	CTCGTTACCT	
A3_7	GACGCAAACG	GTGTTTATTC	TGAACCCCGC	CCTATTGGCA	CTCGTTACCT	
A3_3	GACGCAAACG	GTGTTTATTC	TGAACCCCGC	CCTATTGGCA	CTCGTTACCT	
42_12	AATACTGAGG	GTACTTATTC	AGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
AAV1	GACAACAATG	GACTTTATAC	TGAGCCTCGC	CCCATTGGCA	CCCGTTACCT	
AAV2	GATACTAATG	GCGTGTATTC	AGAGCCTCGC	CCCATTGGCA	CCAGATACCT	
AAV3	GACACTAATG	GTGTTTATAG	TGAACCTCGC	CCTATTGGAA	CCCGGTATCT	
AAV8	AATACAGAAG	GCGTGTACTC	TGAACCCCGC	CCCATTGGCA	CCCGTTACCT	
AAV9	AATACCGAAG	GTGTTTACTC	TGAGCCTCGC	CCCATTGGTA	CTCGTTACCT	
AAV7	GACAGCCAGG	GTGTTTACTC	TGAGCCTCGC	CCTATTGGCA	CTCGTTACCT	
44_2	AACACAGATG	GCACTTATTC	TGAGCCTCGC	CCCATCGGCA	CCCGTTACCT	

Fig. 1AAAL

	4451					4500
		<u>VP1-3 stop</u>		<u>Poly A signal</u>		
	42_2	CACCCGTAAC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	42_8	CACCCGTAAC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGC	TAATTCGTTT
	42_15	CACCCGTAAC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	42_5b	CACCCGTAAC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	42_1b	CACCCGTAAC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TGATTCGTTT
	42_13	CACCCGTAGC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TGATTCGTTT
	42_3a	CACCCGTAAC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	42_4	CACCCGTAAC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	42_5a	CACCCGTAAC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	42_10	CACCCGTAAC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	42_3b	CACCCGTAAC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	42_11	CACCCGTAAC	CTGTAATTAC	TTGTTAATCA	ATAAACCGGT	TGATTCGTTT
	42_6b	CACCCGTAAC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	43_1	CACCCGTAAT	CTGTAATTGC	TTGTTAATCA	ATAAACCGGT	.....
	43_5	CACCCGTAAT	CTGTAATTGC	TTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	43_12	CACCCGTAAT	CTGTAATTGC	TTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	43_20	CACCCGCAAC	CTGTAATTAC	ATGTTAATCA	ATAAACCGGT	TAATTCGTTT
	43_21	CACCCGCAAC	CTGTAATTAC	ATGTTAATCA	ATAAACCGGT	TAATTCGTTT
	43_23	CACCCGCAAC	CTGTAATTAC	ATGTTAATCA	ATAAACCGGT	TAATTCGTTT
	43_25	CACCCGCAAC	CTGTAATTAC	ATGTTAATCA	ATAAACCGGT	TAATTCGTTT
	44_1	CACCCGTAAT	CTGTAATTGC	TCGTTAATCA	ATAAACCGGT	TGATTCGTTT
	44_5	CACCCGTAAT	CTGTAATTGC	TTGTTAATCA	ATAAACCGGT	TGATTCGTTT
	223_10	.....	.....	.....	.....	.....
	223_2	.....	.....	.....	.....	.....
	223_4	.....	.....	.....	.....	.....
	223_5	.....	.....	.....	.....	.....
	223_6	.....	.....	.....	.....	.....
	223_7	.....	.....	.....	.....	.....
	A3_4	TACCCGGAAC	TTGTAATTTT	CTGTTAATGA	ATAAACCGAT	TTATGCGTTT
	A3_5	TACCCGGAAC	TTGTAATTTT	CTGTTAATGA	ATAAACCGAT	TTATGCGTTT
	A3_7	TACCCGGAAC	TTGTAATTTT	CTGTTAATGA	ATAAACCGAT	TTATGCGTTT
	A3_3	TACCCGGAAC	TTGTAATTTT	CTGTTAATGA	ATAAGCCGAT	TTATGCGTTT
	42_12	CACCCGTAAC	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	AAV1	TACCCGTCCC	CTGTAATTAC	GTGTTAATCA	ATAAACCGGT	TGATTCGTTT
	AAV2	GACTCGTAAT	CTGTAATTGC	TTGTTAATCA	ATAAACCGTT	TAATTCGTTT
	AAV3	CACACGAAAC	TTGTGAATCC	TGGTTAATCA	ATAAACCGTT	TAATTCGTTT
	AAV8	CACCCGTAAT	CTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TGATTCGTTT
	AAV9	CACCCGTAAT	TTGTAATTGC	CTGTTAATCA	ATAAACCGGT	TAATTCGTTT
	AAV7	CACCCGTAAT	CTGTAATTGC	ATGTTAATCA	ATAAACCGGT	TGATTCGTTT
	44_2	CACCCGTAAT	CTGTAATTGC	TTGTTAATCA	ATAAACCGGT	TGATTCGTTT
			<u>vp1-3 stop</u>		<u>PolyA signal</u>	

Fig. 1AAAM

4501	42_2 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 42_8 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 42_15 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 42_5b CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTCGTTTA AACCTGCAGG 42_1b CAGTTGAACT TTGGTCTC.. ...AAGGGCG AATTC..... 42_13 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 42_3a CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 42_4 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 42_5a CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 42_10 CAGTTGAACT TTGGTC.... ...AAGGGCG AATTC..... 42_3b CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 42_11 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 42_6b CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 43_1 ..... 43_5 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTCGTTTA AACCTGCAGG 43_12 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 43_20 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 43_21 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 43_23 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 43_25 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 44_1 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 44_5 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... 223_10 ..... 223_2 ..... 223_4 ..... 223_5 ..... 223_6 ..... 223_7 ..... A3_4 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTCGC.GG CCGCTA.... A3_5 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... A3_7 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... A3_3 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTCGT.TT AAACCT.... 42_12 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... AAV1 CAGTTGAACT TTGGTCTCCT GTCCTTCTTA TCTTATCGGT TACCATGGTT AAV2 CAGTTGAACT TTGGTCTC.T GCGTATTTCT ..TTCTT.AT CTAGTTTCCA AAV3 CAGTTGAACT TTGGCTCT.T GTGCACTTCT TTATCTTTAT CTTGTTTCCA AAV8 CAGTTGAACT TTGGTCTC.T GCG..... AAV9 CAGTTGAACT TTGGTCTC.T GCG..... AAV7 CAGTTGAACT TTGGTCTCCT GTGCTTCTTA TCTTATCGGT TTCCATAGCA 44_2 CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC.....	4550
------	---	------

Fig. 1AAAN

	4551				4600
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	ACTAGTCCCT	TTAGTGAGGG	TTAATTCTGA	G.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	AC.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	ATAGCTTACA	CATTAAGTGC	TTGGTTGCGC	T.....	.....
AAV2	TGGCTAC...	GTAAGATAAGT	AGC.....	.....	.....
AAV3	TGGCTACTGC	GTAAGATAAGC	AGCGGCCTGC	GGCGCTTGCG	CTTCGCGGTT
AAV8	.....	.....	.....	.....	.....
AAV9	.....	.....	.....	.....	.....
AAV7	ACTGGTTACA	CATTAAGTGC	TTGGGTGCGC	TTCACGATAA	GAACACTGAC
44_2	.....	.....	.....	.....	.....

Fig. 1AAAO

	4601				4650
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	....CTTGGC	GTAATCATGG	GTCATAG...	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	....TCGCGA	TAAAAGACTT	ACGTCATCGG	GTTACCCCTA	GTGATGGAGT
AAV2	....ATGGCG	GGTTAATCAT	TAACTACAAG	GA.ACCCCTA	GTGATGGAGT
AAV3	TACAACTGCT	GGTTAATATT	TAACTCTCGC	CATACCTCTA	GTGATGGAGT
AAV8	.....	.....	.....	.....	.....
AAV9	.....	.....	.....	.....	.....
AAV7	.....	.....	..GTCACCGC	GGTACCCCTA	GTGATGGAGT
44_2	.....	.....	.....	.....	.....

Fig. 1AAAP

	4651				4700
42_2	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....
AAV1	TGCCCCACTCC	CTCTCTGCGC	GCTCGCTCGC	TCGGTGGGGC	CTGCCGACCA
AAV2	TGGCCACTCC	CTCTCTGCGC	GCTCGCTCGC	TCACTGAGGC	CGGGCGACCA
AAV3	TGGCCACTCC	CTCTATGCGC	ACTCGCTCGC	TCGGTGGGGC	CTGGCGACCA
AAV8	.....	.....	.....	.....	.....
AAV9	.....	.....	.....	.....	.....
AAV7	TGGCCACTCC	CTCTATGCGC	GCTCGCTCGC	TCGGTGGGGC	CTGCCGACCA
44_2	.....	.....	.....	.....	.....

Fig. 1AAAQ

	4701					4750
42_2	.....	.....	.....	.....	.....	.....
42_8	.....	.....	.....	.....	.....	.....
42_15	.....	.....	.....	.....	.....	.....
42_5b	.....	.....	.....	.....	.....	.....
42_1b	.....	.....	.....	.....	.....	.....
42_13	.....	.....	.....	.....	.....	.....
42_3a	.....	.....	.....	.....	.....	.....
42_4	.....	.....	.....	.....	.....	.....
42_5a	.....	.....	.....	.....	.....	.....
42_10	.....	.....	.....	.....	.....	.....
42_3b	.....	.....	.....	.....	.....	.....
42_11	.....	.....	.....	.....	.....	.....
42_6b	.....	.....	.....	.....	.....	.....
43_1	.....	.....	.....	.....	.....	.....
43_5	.....	.....	.....	.....	.....	.....
43_12	.....	.....	.....	.....	.....	.....
43_20	.....	.....	.....	.....	.....	.....
43_21	.....	.....	.....	.....	.....	.....
43_23	.....	.....	.....	.....	.....	.....
43_25	.....	.....	.....	.....	.....	.....
44_1	.....	.....	.....	.....	.....	.....
44_5	.....	.....	.....	.....	.....	.....
223_10	.....	.....	.....	.....	.....	.....
223_2	.....	.....	.....	.....	.....	.....
223_4	.....	.....	.....	.....	.....	.....
223_5	.....	.....	.....	.....	.....	.....
223_6	.....	.....	.....	.....	.....	.....
223_7	.....	.....	.....	.....	.....	.....
A3_4	.....	.....	.....	.....	.....	.....
A3_5	.....	.....	.....	.....	.....	.....
A3_7	.....	.....	.....	.....	.....	.....
A3_3	.....	.....	.....	.....	.....	.....
42_12	.....	.....	.....	.....	.....	.....
AAV1	AAGGTCCGCA	GACGGCAGAG	CTCTGCTCTG	CCGGCCCCAC	CGAGCGAGCG	
AAV2	AAGGTCGCCC	GACGCCCGGG	CTTTGCCCGG	GCGGCCTCAG	TGAGCGAGCG	
AAV3	AAGGTCGCCA	GACGGACGTG	CTTTGCACGT	CCGGCCCCAC	CGAGCGAGCG	
AAV8	.....	.....	.....	.....	.....	.....
AAV9	.....	.....	.....	.....	.....	.....
AAV7	AAGGTCCGCA	GACGGCAGAG	CTCTGCTCTG	CCGGCCCCAC	CGAGCGAGCG	
44_2	.....	.....	.....	.....	.....	.....

Fig. 1AAAR

	4751		4774
42_2	.....	.....	.....
42_8	.....	.....	.....
42_15	.....	.....	.....
42_5b	.....	.....	.....
42_1b	.....	.....	.....
42_13	.....	.....	.....
42_3a	.....	.....	.....
42_4	.....	.....	.....
42_5a	.....	.....	.....
42_10	.....	.....	.....
42_3b	.....	.....	.....
42_11	.....	.....	.....
42_6b	.....	.....	.....
43_1	.....	.....	.....
43_5	.....	.....	.....
43_12	.....	.....	.....
43_20	.....	.....	.....
43_21	.....	.....	.....
43_23	.....	.....	.....
43_25	.....	.....	.....
44_1	.....	.....	.....
44_5	.....	.....	.....
223_10	.....	.....	.....
223_2	.....	.....	.....
223_4	.....	.....	.....
223_5	.....	.....	.....
223_6	.....	.....	.....
223_7	.....	.....	.....
A3_4	.....	.....	.....
A3_5	.....	.....	.....
A3_7	.....	.....	.....
A3_3	.....	.....	.....
42_12	.....	.....	.....
AAV1	AGCGCGCAGA	GAGGGAGTGG	GCAA
AAV2	AGCGCGCAGA	GAGGGAGTGG	CCAA
AAV3	AGTGCGCATA	GAGGGAGTGG	CCAA
AAV8	.....	.....	.....
AAV9	.....	.....	.....
AAV7	AGCGCGCATA	GAGGGAGTGG	CCAA
44_2	.....	.....	.....

Fig. 2A

149 CLVP1 MAADGTLPMLENDLSIGIREWMLKFGAPKPFANQKQDDGRGLVPGTYLGPFGNGLKGEFVNADAAALEHDKAYDQQLKAGDNPTLYRNHADAEFOERLQEDTSFGNIGRAVFOAKKRVLEPLGLVEEGANTAPGKRPVESP  
 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

Fig. 2B

[illegible]

Fig. 2C

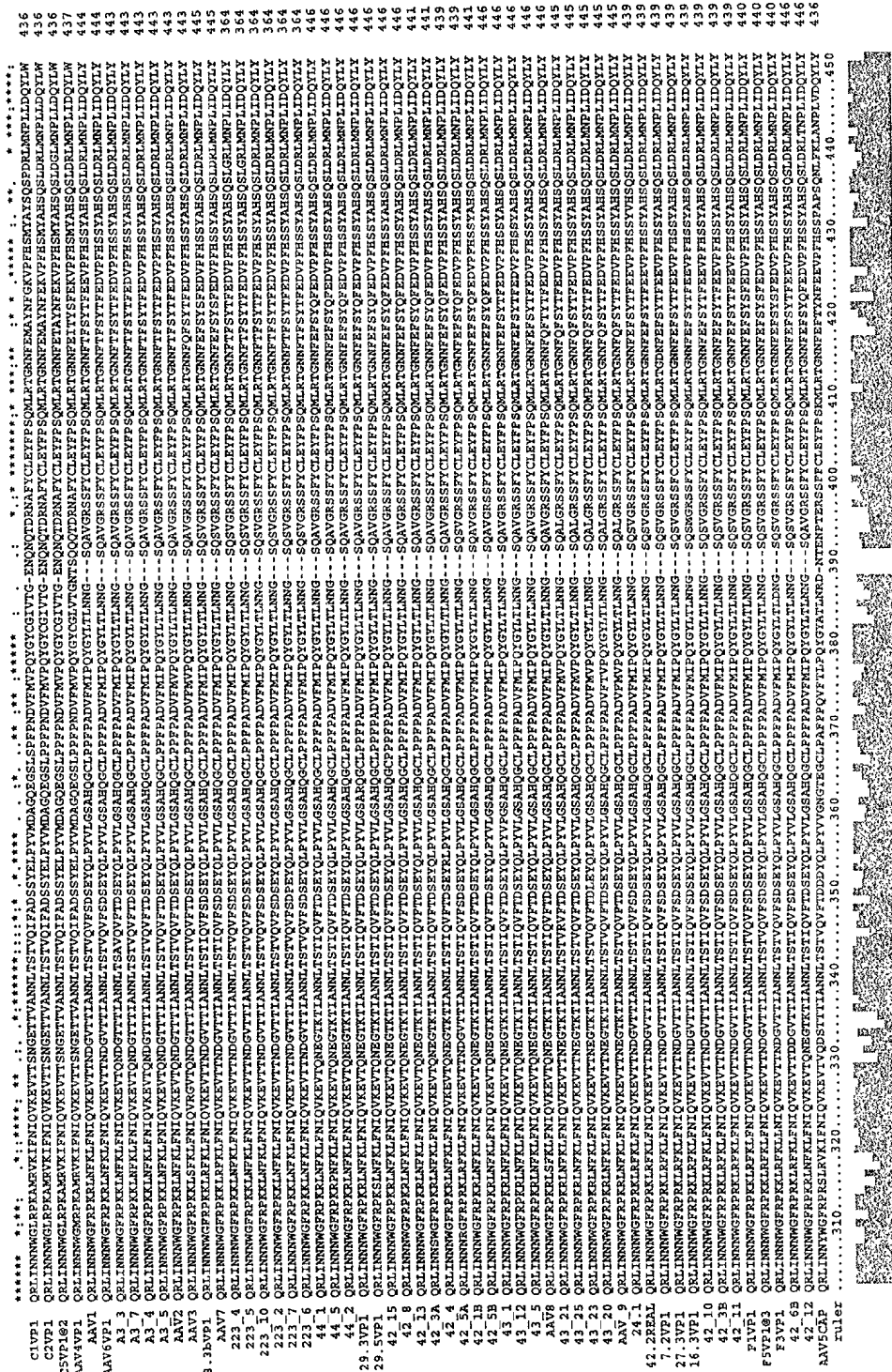


Fig. 2D

[illegible]

Fig. 2E

[illegible]

Fig. 2F

C1VP1 L 733  
 C2VP1 L 733  
 C5VP102 L 733  
 RAV4VP1 L 734  
 RAV1 L 736  
 RAV6VP1 L 736  
 A3 3 L 735  
 A3 7 L 735  
 A3 4 L 735  
 A3 5 L 735  
 RAV2 L 735  
 RAV3 L 736  
 3.3bVP1 L 737  
 RAV1 L 737  
 223 4 - 644  
 223 5 - 644  
 223 10 - 644  
 223 2 - 644  
 223 7 - 644  
 223 6 - 644  
 44 1 L 738  
 44 5 L 738  
 44 2 L 738  
 29.3bVP1 L 738  
 29.5bVP1 L 738  
 42 15 L 738  
 42 8 L 738  
 42 13 L 733  
 42 38 L 733  
 42 4 L 731  
 42 58 L 731  
 42 18 L 733  
 42 58 L 738  
 43 1 L 738  
 43 12 L 738  
 43 5 L 738  
 RAV8 L 738  
 43 21 L 736  
 43 25 L 736  
 43 23 L 736  
 43 20 L 736  
 RAV 9 L 736  
 24 1 L 728  
 42.2bRAL L 728  
 7.2VP1 L 728  
 27.3VP1 L 728  
 16.3VP1 L 728  
 42 10 L 728  
 42 38 L 728  
 42 11 L 728  
 RAV1 L 729  
 P5VP103 L 729  
 P3VP1 L 729  
 42 68 L 735  
 42 12 L 735  
 RAV5VP1 L 724  
 ruler .

Fig. 3A

Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Val Gln Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Leu His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Val Gln Thr Ile Tyr Arg Gly Val Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Glu Tyr Ile  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Ala Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Leu Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220

Fig. 3B

Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ala Leu Thr Lys Ser Ala Pro Asp Tyr Leu Val Gly Pro Ser  
 275 280 285  
 Leu Pro Ala Asp Ile Lys Thr Asn Arg Ile Tyr Arg Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Ala Tyr Ala Gly Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Gln Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445

Fig. 3C

Glu Leu Thr Arg Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460

Glu Val Lys Glu Phe Phe Arg Trp Ala Ser Asp His Val Thr Glu Val  
 465 470 475 480

Ala His Glu Phe Tyr Val Arg Lys Gly Gly Ala Ser Lys Arg Pro Ala  
 485 490 495

Pro Asp Asp Ala Asp Ile Ser Glu Pro Lys Arg Ala Cys Pro Ser Val  
 500 505 510

Ala Asp Pro Ser Thr Ser Asp Ala Glu Gly Ala Pro Val Asp Phe Ala  
 515 520 525

Asp Arg Tyr Gln Asn Lys Cys Ser Arg His Ala Gly Met Ile Gln Met  
 530 535 540

Leu Phe Pro Cys Lys Thr Cys Glu Arg Met Asn Gln Asn Phe Asn Ile  
 545 550 555 560

Cys Phe Thr His Gly Val Arg Asp Cys Leu Glu Cys Phe Pro Gly Val  
 565 570 575

Ser Glu Ser Gln Pro Val Val Arg Lys Lys Thr Tyr Arg Lys Leu Cys  
 580 585 590

Ala Ile His His Leu Leu Gly Arg Ala Pro Glu Ile Ala Cys Ser Ala  
 595 600 605

Cys Asp Leu Val Asn Val Asp Leu Asp Asp Cys Val Ser Glu Gln  
 610 615 620